

PCTWORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁵ : C12S 3/22, C12N 9/40, 15/10, 15/56, 15/63, A01N 1/02	A1	(11) International Publication Number: WO 94/23070 (43) International Publication Date: 13 October 1994 (13.10.94)
(21) International Application Number: PCT/US94/03338 (22) International Filing Date: 28 March 1994 (28.03.94) (30) Priority Data: 08/037,248 26 March 1993 (26.03.93) US (71) Applicant: NEW YORK BLOOD CENTER, INC. [US/US]; 310 East 67th Street, New York, NY 10021 (US). (72) Inventors: ZHU, Alex; 155 East 31st Street, Apartment 14E, New York, NY 10016 (US). GOLDSTEIN, Jack; 450 East 63rd Street, New York, NY 10021 (US). (74) Agents: GEORGE, Kenneth et al.; Amster, Rothstein & Ebenstein, 90 Park Avenue, New York, NY 10016 (US).		(81) Designated States: AU, CA, CN, JP, NZ, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE). Published <i>With international search report.</i>
(54) Title: RECOMBINANT α -N-ACETYL GALACTOSAMINIDASE ENZYME AND cDNA ENCODING SAID ENZYME (57) Abstract This invention relates to a recombinant enzyme for use in the removal of A antigens from the surface of cells in blood products. Specifically, this invention is directed to a recombinant α -N-acetylgalactosaminidase enzyme from chicken liver, methods of cloning and expressing said recombinant α -N-acetylgalactosaminidase enzyme and a method of removing A antigens from the surface of cells in blood products using said recombinant α -N-acetylgalactosaminidase enzyme.		

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

RECOMBINANT α -N-ACETYLGALACTOSAMINIDASE
ENZYME AND cDNA ENCODING SAID ENZYME

Statement of Government Interest

This invention was made with government support under
5 NMRDC Grant Number N0014-90-J-1638. As such, the government
has certain rights in the invention.

CROSS REFERENCE TO RELATED APPLICATIONS

This Application is a Continuation-In-Part of
Application Serial No. 07/964,756 filed October 22, 1992,
10 entitled PREPARATION OF ENZYME FOR CONVERSION OF SUB-TYPE A AND
AB ERYTHROCYTES.

FIELD OF THE INVENTION

This invention relates to a recombinant enzyme for use
in the removal of type A antigens from the surface of cells in
15 blood products, thereby converting certain sub-type A blood
products to type O blood products and certain type AB blood
products to type B blood products. This invention further
relates to methods of cloning and expressing said recombinant
enzyme. More particularly, this invention is directed to a
20 recombinant chicken liver α -N-acetylgalactosaminidase enzyme,
methods of cloning and expressing said recombinant
 α -N-acetylgalactosaminidase enzyme, and a method of removing
type A antigens from the surface of cells in type A and AB

blood products using said recombinant α -N-acetylgalactosaminidase enzyme by contacting said enzyme with blood products so as to remove the terminal moiety of the A-antigenic determinant from the surface of cells (for example, erythrocytes) in said blood products, while allowing the structure and function of the cells in the blood products to remain intact. The recombinant α -N-acetylgalactosaminidase enzyme of this invention provides a readily available and cost-efficient enzyme which can be used in the removal of type A antigens from the surface of cells in type A and AB blood products. Treatment of certain sub-type A blood products with the recombinant enzyme of this invention provides a source of cells free of the A antigen, which blood products are thereby rendered useful in transfusion therapy in the same manner of O type blood products.

BACKGROUND OF THE INVENTION

As used herein, the term "blood products" includes whole blood and cellular components derived from blood, including erythrocytes (red blood cells) and platelets.

There are more than thirty blood group (or type) systems, one of the most important of which is the ABO system. This system is based on the presence or absence of antigens A and/or B. These antigens are found on the surface of erythrocytes and on the surface of all endothelial and most

epithelial cells as well. The major blood product used for transfusion is erythrocytes, which are red blood cells containing hemoglobin, the principal function of which is the transport of oxygen. Blood of group A contains antigen A on its erythrocytes. Similarly, blood of group B contains antigen B on its erythrocytes. Blood of group AB contains both antigens, and blood of group O contains neither antigen.

The blood group structures are glycoproteins or glycolipids and considerable work has been done to identify the specific structures making up the A and B determinants or antigens. It has been found that the blood group specificity is determined by the nature and linkage of monosaccharides at the ends of the carbohydrate chains. The carbohydrate chains are attached to a peptide or lipid backbone which is embedded in the lipid bi-layer of the membrane of the cells. The most important (immuno-dominant or immuno-determinant) sugar has been found to be N-acetylgalactosamine for the type A antigen and galactose for the type B antigen.

There are three recognized major sub-types of blood type A. These sub-types are known as A_1 , A intermediate (A_{int}) and A_2 . There are both quantitative and qualitative differences which distinguish these three sub-types. Quantitatively, A_1 erythrocytes have more antigenic A sites, i.e., terminal N-acetylgalactosamine residues, than A_{int} erythrocytes which in turn have more antigenic A sites than

A₂ erythrocytes. Qualitatively, the transferase enzymes responsible for the formation of A antigens differ biochemically from each other in A₁, A_{int} and A₂ individuals. Some A antigens found in A₁ cells contain dual
5 A antigenic sites.

Blood of group A contains antibodies to antigen B. Conversely, blood of group B contains antibodies to antigen A. Blood of group AB has neither antibody, and blood group O has both. A person whose blood contains either (or both) of the
10 anti-A or anti-B antibodies cannot receive a transfusion of blood containing the corresponding incompatible antigen(s). If a person receives a transfusion of blood of an incompatible group, the blood transfusion recipient's antibodies coat the red blood cells of the transfused incompatible group and cause
15 the transfused red blood cells to agglutinate, or stick together. Transfusion reactions and/or hemolysis (the destruction of red blood cells) may result therefrom.

In order to avoid red blood cell agglutination, transfusion reactions and hemolysis, transfusion blood type is
20 cross-matched against the blood type of the transfusion recipient. For example, a blood type A recipient can be safely transfused with type A blood which contains compatible antigens. Because type O blood contains no A or B antigens, it can be transfused into any recipient with any blood type, i.e.,
25 recipients with blood types A, B, AB or O. Thus, type O blood

is considered "universal", and may be used for all transfusions. Hence, it is desirable for blood banks to maintain large quantities of type O blood. However, there is a paucity of blood type O donors. Therefore, it is useful to
5 convert types A, B and AB blood to type O blood in order to maintain large quantities of universal blood products.

In an attempt to increase the supply of type O blood, methods have been developed for converting certain type A, B and AB blood to type O blood. For example, U.S. Patent No.
10 4,609,627 entitled "Enzymatic Conversion of Certain Sub-Type A and AB Erythrocytes" ("the '627 Patent"), which is incorporated herein by reference, is directed to a process for converting A_{int} and A_2 (including A_2B erythrocytes) to erythrocytes of the H antigen type, as well as to compositions of type B
15 erythrocytes which lack A antigens, which compositions, prior to treatment, contained both A and B antigens on the surface of said erythrocytes. The process for converting A_{int} and A_2 erythrocytes to erythrocytes of the H antigen type which is described in the '627 Patent includes the steps of
20 equilibrating certain sub-type A or AB erythrocytes, contacting the equilibrated erythrocytes with purified chicken liver α -N-acetylgalactosaminidase enzyme for a period sufficient to convert the A antigen to the H antigen, removing the enzyme from the erythrocytes and re-equilibrating the erythrocytes.
25 As described in the '627 Patent, α -N-acetylgalactosaminidase

obtained from an avian liver (specifically, chicken liver) source was found to have superior activity in respect of enzymatic conversion or cleavage of A antigenic sites.

Prior to the present invention, it was necessary to
5 purify the enzyme from an avian liver source, a process which is time consuming and can be expensive. Hence, a need has arisen to develop an enzyme source which is more readily available. In addition, a need has arisen to develop an enzyme useful in blood product conversion which enzyme is
10 cost-efficient.

A simplified purification process is described in a related application, Serial No. 07/964,756, filed October 22, 1992, entitled "Preparation of Enzyme for Conversion of Sub-Type A and AB Erythrocytes". This process, as described in
15 the related application, utilizes chicken liver as a source of enzyme and, therefore, requires a number of purification steps. Despite this simplified process, it is still desirable to provide a more readily available and controlled source of enzyme, that being cloned and expressed enzyme. This would
20 provide an enzyme source which is more consistent and which is readily purified at less cost and expense, with a still further reduced number of purification steps. Additionally, a recombinant, cloned enzyme allows for specific protein sequence modifications, which can be introduced to generate an enzyme
25

with optimized specific activity, substrate specificity and pH range.

α -N-acetylgalactosaminidase enzymes are characterized (and thereby named) by their ability to cleave N-acetylgalactosamine sugar groups. In isolating or identifying these enzymes, their activity is assessed in the laboratory by evaluating cleavage of synthetic substrates which mimic the sugar groups cleaved by the enzymes, with p-nitrophenylglycopyranoside derivatives of the target sugar groups being commonly used. Although very useful in enzyme identification and isolation procedures (the quantitative cleavage of these synthetic substrates can be used to readily distinguish (and thereby identify) enzymes isolated from different sources), these synthetic substrates are simple structurally and small-sized and mimic only a portion of the natural glycoproteins and glycolipid structures which are of primary concern, those being the A antigens on the surface of cells.

A natural glycolipid substrate, originally isolated from sheep erythrocytes, is the Forsmann antigen (globopentaglycosylceramide). The Forsmann antigen substrate appropriately mimics the natural A antigen glycolipid structures and is therefore utilized to predict the activity of α -N-acetylgalactosaminidase enzymes against the A antigen substrate. Isolated Forsmann antigen glycolipids have been

shown to inhibit hemolysis of sheep red cells by immune rabbit anti-A serum in the presence of serum complement.

α -N-acetylgalactosaminidase enzyme has been isolated from a number of sources besides chicken liver (described above), including bacteria, mollusks, earthworms, and human liver. The human α -N-acetylgalactosaminidase enzyme has been purified, sequenced, cloned and expressed. For example, in "Human α -N-Acetylgalactosaminidase - Molecular Cloning, Nucleotide Sequence and Expression of a Full-length cDNA", by Wang et al., in The Journal of Biological Chemistry, Vol. 265, No. 35, pages 21859-21866 (December 15, 1990), the cDNA encoding human α -N-acetylgalactosaminidase was sequenced. In addition, in "Molecular Cloning of a Full-Length cDNA for Human α -N-Acetylgalactosaminidase (α -Galactosidase B)", by Tsuji et al., in Biochemical And Biophysical Research Communications, Vol. 163, No. 3, pages 1498-1504 (September 29, 1989), the cDNA encoding human α -N-acetylgalactosaminidase was sequenced. Both the nucleotide sequence and the amino acid sequence of human α -N-acetylgalactosaminidase is published therein. Further, PCT Application No. WO 92/07936 discloses the cloning and expression of the cDNA which encodes human α -N-acetylgalactosaminidase.

Although human α -N-acetylgalactosaminidase has been purified, sequenced, cloned and expressed, it is not appropriate for use in removing A antigens from the surface of

cells in blood products. In determining whether an enzyme is appropriate for use in removing A antigens from the surface of cells, one must consider the following enzyme characteristics, particularly with respect to the Forsmann antigen substrate:

5 substrate specificity, specific activity or velocity of the substrate cleavage reaction, and pH optimum. Substrate specificity is measured in the K_m value, which measures the binding constant or affinity of an enzyme for a particular substrate. The lower a K_m value, the more tightly an enzyme

10 binds its substrate. The velocity of an enzyme cleavage reaction is measured in the V_{max} , the reaction rate at a saturating concentration of substrate. A higher V_{max} indicates a faster cleavage rate. The ratio of these two parameters, V_{max}/K_m , is a measure of the overall efficiency of an enzyme in

15 reacting with (cleaving) a given substrate. A higher V_{max}/K_m indicates greater enzyme efficiency. For successful and clinically applicable removal of A antigens from the surface of cells, the enzyme must be sufficiently active at or above a pH at which the cells being treated can be maintained. The

20 procedure described in the '627 patent calls for treatment of cells at or above a pH of 5.6. Therefore, the pH optimum of an appropriate enzyme must still provide reasonable enzyme activity at this pH. These specific characteristics (V_{max}/K_m , V_{max} , K_m and pH optimum) are reported for the human

25 α -N-acetylgalactosaminidase enzyme in "Studies on Human

Liver α -galactosidases", by Dean et al. in The Journal of Biological Chemistry, Vol. 254, No. 20, pages 10001-10005 (1979).

The V_{max}/K_m value for the Forsmann antigen of human
5 α -N-acetylgalactosaminidase is 0.46, as compared to a V_{max}/K_m
value of 5.0 for the chicken liver enzyme, indicating an
approximately ten-fold difference in efficiency. The K_m is
lower and the V_{max} is higher for the chicken liver enzyme,
compared to the human enzyme. Further, human
10 α -N-acetylgalactosaminidase has a pH optimum for the Forsmann
antigen of 3.9, compared to 4.7 for chicken liver
 α -N-acetylgalactosaminidase. By all of these enzyme
characteristics, human α -N-acetylgalactosaminidase enzyme is
not suitable for removal of A antigens, particularly when
15 compared to the chicken liver enzyme.

As a result, a need still existed to develop an enzyme
which is capable of removing A antigens from the surface of
cells in blood products, wherein said enzyme is readily
available and cost-efficient.

20 It is therefore an object of this invention to provide
a recombinant enzyme for use in the removal of A antigens from
the surface of cells in blood products.

It is another object of this invention to provide a
recombinant enzyme for use in the removal of A antigens from
25 the surface of cells in blood products wherein said enzyme is

readily available and may be manufactured on a cost-efficient basis.

It is a further object of this invention to provide methods of cloning and expressing a recombinant enzyme useful
5 in the removal of A antigens from the surface of cells in blood products.

It is yet another object of this invention to provide a method of removing A antigens from the surface of cells in blood products using a recombinant enzyme.

10

BRIEF DESCRIPTION OF THE DRAWINGS

The above brief description, as well as further objects and features of the present invention, will be more fully understood by reference to the following detailed description of the presently preferred, albeit illustrative,
15 embodiment of the present invention when taken in conjunction with the accompanying drawing wherein:

Figure 1 represents a diagram of the strategy used to clone and sequence the chicken liver α -N-acetylgalactosaminidase cDNA;

20 Figure 2 represents the nucleic acid sequence and the deduced amino acid sequence of the chicken liver α -N-acetylgalactosaminidase cDNA clone;

Figure 3 represents the expression of chicken liver α -N-acetylgalactosaminidase in bacteria and rabbit
25 reticulocyte lysate as shown by Western blot;

Figure 4 represents a homology comparison between α -N-acetylgalactosaminidases and α -galactosidases; and

Figure 5 represents the expression of chicken liver α -N-acetylgalactosaminidase in yeast as shown by Western blot.

5

SUMMARY OF THE INVENTION

This invention is directed to a recombinant chicken liver α -N-acetylgalactosaminidase enzyme, which enzyme has a molecular weight of about 45 kDa, is immunoreactive with an antibody specific for chicken liver α -N-acetylgalactosaminidase, and also has about 80% amino acid sequence homology with human α -N-acetylgalactosaminidase enzyme. The recombinant chicken liver α -N-acetylgalactosaminidase enzyme of this invention has the amino acid sequence depicted in Figure 2, from amino acid number 1 to amino acid number 406. This invention is further directed to methods of cloning and expressing the recombinant chicken liver α -N-acetylgalactosaminidase enzyme, and to a method of using said enzyme to remove A antigens from the surface of cells in blood products so as to convert said blood products of certain A sub-types to type O, thereby rendering said blood products universal for use in transfusion therapy.

DETAILED DESCRIPTION OF THE INVENTION

This invention is directed to a recombinant enzyme for use in the removal of type A antigens from the surface of cells

in blood products, thereby converting certain sub-type A blood products to type O blood products and certain sub-type AB blood products to type B blood products. The recombinant chicken liver α -N-acetylgalactosaminidase enzyme of this invention has a molecular weight of about 45 kDa and is immunoreactive with an antibody specific for chicken liver α -N-acetylgalactosaminidase. In addition, the recombinant enzyme of this invention has about 80% amino acid sequence homology with human α -N-acetylgalactosaminidase enzyme. The recombinant chicken liver α -N-acetylgalactosaminidase enzyme of this invention has the following nucleic acid and deduced amino acid sequence:

SEQ ID NO 1:

15	ATG CTG GAG AAC GGG CTG GCG CGG ACC CCG CCC ATG GGC TGG TTG GCC Met Leu Glu Asn Gly Leu Ala Arg Thr Pro Pro Met Gly Trp Leu Ala
	TGG GAG CGG TTC CGC TGC AAC GTG AAC TGC CGG GAG GAC CCC CGC CAG Trp Glu Arg Phe Arg Cys Asn Val Asn Cys Arg Glu Asp Pro Arg Gln
	TGC ATC AGT GAG ATG CTC TTC ATG GAG ATG GCA GAC CGA ATA GCA GAG Cys Ile Ser Glu Met Leu Phe Met Glu Met Ala Asp Arg Ile Ala Glu
20	GAC GGC TGG AGG GAG CTG GGC TAC AAG TAC ATC AAT ATC GAT GAC TGC Asp Gly Trp Arg Glu Leu Gly Tyr Lys Tyr Ile Asn Ile Asp Asp Cys
	TGG GCC GCC AAG CAG CGT GAC ACT GAG GGG CGG CTG GTG CCT GAC CCC Trp Ala Ala Lys Gln Arg Asp Thr Glu Gly Arg Leu Val Pro Asp Pro
25	GAG AGG TTC CCC CGG GGC ATT AAG GCC TTG GCT GAC TAC GTT CAT GCC Glu Arg Phe Pro Arg Gly Ile Lys Ala Leu Ala Asp Tyr Val His Ala
	CGA GGC TTG AAG CTG GGC ATT TAT GGC GAC CTG GGC AGA CTC ACC TGT Arg Gly Leu Lys Leu Gly Ile Tyr Gly Asp Leu Gly Arg Leu Thr Cys
	GGA GGC TAC CCA GGC ACC ACG CTG GAC CGT GTG GAG CAG GAC GCA CAG Gly Gly Tyr Pro Gly Thr Thr Leu Asp Arg Val Glu Gln Asp Ala Gln

	ACC	TTC	GCT	GAG	TGG	GGT	GTG	GAC	ATG	CTG	AAG	CTA	GAT	GGG	TGC	TAC
	Thr	Phe	Ala	Glu	Trp	Gly	Val	Asp	Met	Leu	Lys	Leu	Asp	Gly	Cys	Tyr
	TCA	TCG	GGG	AAG	GAG	CAG	GCA	CAG	GGC	TAC	CCA	CAA	ATG	GCA	AGG	GCC
	Ser	Ser	Gly	Lys	Glu	Gln	Ala	Gln	Gly	Tyr	Pro	Gln	Met	Ala	Arg	Ala
5	TTG	AAC	GCC	ACT	GGC	CGC	CCC	ATC	GTC	TAC	TCC	TGC	AGC	TGG	CCA	GCC
	Leu	Asn	Ala	Thr	Gly	Arg	Pro	Ile	Val	Tyr	Ser	Cys	Ser	Trp	Pro	Ala
	TAC	CAG	GGG	GGG	CTG	CCT	CCC	AAG	GTG	AAC	TAC	ACT	CTC	CTG	GGT	GAG
	Tyr	Gln	Gly	Gly	Leu	Pro	Pro	Lys	Val	Asn	Tyr	Thr	Leu	Leu	Gly	Glu
10	ATC	TGC	AAC	CTG	TGG	CGG	AAC	TAC	GAT	GAC	ATC	CAG	GAC	TCA	TGG	GAC
	Ile	Cys	Asn	Leu	Trp	Arg	Asn	Tyr	Asp	Asp	Ile	Gln	Asp	Ser	Trp	Asp
	AGC	GTG	CTT	TCC	ATC	GTG	GAC	TGG	TTC	TTC	ACA	AAC	CAG	GAT	GTG	CTG
	Ser	Val	Leu	Ser	Ile	Val	Asp	Trp	Phe	Phe	Thr	Asn	Gln	Asp	Val	Leu
	CAG	CCG	TTT	GCT	GGC	CCT	GGC	CAC	TGG	AAT	GAC	CCA	GAC	ATG	CTC	ATC
	Gln	Pro	Phe	Ala	Gly	Pro	Gly	His	Trp	Asn	Asp	Pro	Asp	Met	Leu	Ile
15	ATT	GGA	AAT	TTC	GGT	CTC	AGC	TAT	GAG	CAG	TCA	CGT	TCC	CAA	ATG	GCC
	Ile	Gly	Asn	Phe	Gly	Leu	Ser	Tyr	Glu	Gln	Ser	Arg	Ser	Gln	Met	Ala
	TTG	TGG	ACC	ATT	ATG	GCA	GCT	CCA	CTC	CTC	ATG	TCC	ACC	GAC	CTG	CGC
	Leu	Trp	Thr	Ile	Met	Ala	Ala	Pro	Leu	Leu	Met	Ser	Thr	Asp	Leu	Arg
20	ACT	ATC	TCG	CCG	AGT	GCC	AAG	AAG	ATT	CTG	CAG	AAC	CGC	CTG	ATG	ATC
	Thr	Ile	Ser	Pro	Ser	Ala	Lys	Lys	Ile	Leu	Gln	Asn	Arg	Leu	Met	Ile
	CAG	ATA	AAC	CAG	GAC	CCC	TTG	GGA	ATC	CAG	GGG	CGC	AGG	ATC	ATC	AAG
	Gln	Ile	Asn	Gln	Asp	Pro	Leu	Gly	Ile	Gln	Gly	Arg	Arg	Ile	Ile	Lys
	GAG	GGA	TCC	CAC	ATT	GAG	GTG	TTC	CTG	CGC	CCG	CTG	TCA	CAG	GCT	GCC
	Glu	Gly	Ser	His	Ile	Glu	Val	Phe	Leu	Arg	Pro	Leu	Ser	Gln	Ala	Ala
25	AGT	GCC	CTG	GTC	TTC	TTC	AGC	CGG	AGG	ACA	GAC	ATG	CCC	TTC	CGC	TAC
	Ser	Ala	Leu	Val	Phe	Phe	Ser	Arg	Arg	Thr	Asp	Met	Pro	Phe	Arg	Tyr
	ACC	ACC	AGT	CTT	GCC	AAG	CTT	GGC	TTC	CCC	ATG	GGA	GCT	GCA	TAT	GAG
	Thr	Thr	Ser	Leu	Ala	Lys	Leu	Gly	Phe	Pro	Met	Gly	Ala	Ala	Tyr	Glu
30	GTG	CAA	GAC	GTG	TAC	AGT	GGG	AAG	ATC	ATC	AGT	GGC	CTG	AAG	ACA	GGA
	Val	Gln	Asp	Val	Tyr	Ser	Gly	Lys	Ile	Ile	Ser	Gly	Leu	Lys	Thr	Gly
	GAC	AAC	TTC	ACA	GTG	ATC	ATC	AAC	CCC	TCA	GGG	GTG	GTG	ATG	TGG	TAC
	Asp	Asn	Phe	Thr	Val	Ile	Ile	Asn	Pro	Ser	Gly	Val	Val	Met	Trp	Tyr
	CTG	TGT	CCC	AAA	GCA	CTG	CTC	ATC	CAG	CAG	CAA	GCT	CCT	GGG	GGG	CCC
	Leu	Cys	Pro	Lys	Ala	Leu	Leu	Ile	Gln	Gln	Gln	Ala	Pro	Gly	Gly	Pro


```

TCG CGC CTG CCC CTT CTG TGA GGC CCA TGA TTG GGA GCC CTG GGA TAC
Ser Arg Leu Pro Leu Leu ***

ATC TCA CCG CTG CTC AAG TGC CTT CTT CTG GTG TGG CTG GGG GAG GAC
ATG CAG CTT GCT CCT CTG GCA CCA CCT GAT GAT TTC TAC TCA TTC CAC
5 GTG AAG CAG GAC TTC TTG TTA CTC CCT CCT GAG AGC ATG CAA AGC GCT
CTG AGG TCC TCC TGT GGA AGA GGA GTG TTC CCA GTG ACC ATC CTT TAG
GAC CAG ATG TGG TCA CCT TTT TTC CTT TGC TTG GCT TAG GAC AAA GGG
CTG TCC ACA GGC TGC ACC CCT CTT CCC AGG CAC CAT CCC CAG ACC AGG
AGC TCC TGG GGC CAG GCT GTC TCT GTC TGG CAG CAG GAT CAG CAG GTA
10 ACA CCA CTA CAG TGT AGT CCG CAC ATA ATG AAA AAG AAA TCT AAA CAA
AAC GTG TGC CAG TAG TGT ACT GAA CCC GCT CTG GTT ACA GCA GAG CAA
AAC CTG AGT TGT CCA TGC ACA ATC CCA GTA TCC TCA CTG TGG TGT TAG
CAT GAA AAA TTG CAG TCA CAG TGC ATT GTG CAC GAG TGG TGT CTG GAA
GAT GCT GAT GCT TGT TCG TGG TGG TCT TAA GGT GGG AGA TGC TCA TGG
15 GTG CTG GCC AAG TTG CAT CTC AAT CTT GTG AGG CTG AAC CTT CCA GCA
TTT CTC AGG GAA AGG CTC TTC CTT TTA AAG GCA GCC TGC ACA AAT AGA
AGG GGC TCA GAA GGA CGC ACG AGG AGG GGC TCA GGT GGG CCG TGC TCC
CCT GAC CAC CCC AAG AGG GGT CAA CTA CTC ACC AAA ATC TAC CCC TTT
CAA GGC CAG GTC AGC CCA GGG AGA CGC ACC CAA GGT TAA ACC TCA AAA
20 CAG GAA ATC ACC CTA TTT TAA ATT AGT GAG AAA TTG AAC TTC CCC ATT
CTA TTC AGA TGA GGG CTA GAA GCC CAC TCT CCT TAG AAG GCA CGT GGT
GGA TTC CTG CCC CTT GCA GAG ACA TTG TGG TCT GAA GCA AGA TGC TGA
ATG TGA TCT TTG CAG CGC TGG AAA TGA CAT GTC TGT TTC ATG CTT GTG
TGG GAG ATG GCT TTG TTT TTG TGA TTT TGA CAA TTT AAC TGA AAT AAA
25 AGG GAA GCA GAG GGG

```

A DNA vector containing a sequence encoding chicken liver α -N-acetylgalactosaminidase was deposited under the Budapest Treaty with the American Type Culture Collection, Rockville, Maryland, on March 17, 1993, tested and found viable
5 on March 22, 1993 and catalogued as ATCC No. 75434.

The recombinant chicken liver α -N-acetylgalactosaminidase enzyme of this invention can be cloned and expressed so that it is readily available for use in the removal of A antigens from the surface of cells in blood
10 products. The enzyme of this invention can be cloned and expressed by screening a chicken liver cDNA library to obtain the cDNA sequence which encodes the chicken liver α -N-acetylgalactosaminidase, sequencing the encoding cDNA once it is determined, cloning the encoding cDNA and expressing
15 α -N-acetylgalactosaminidase from the cloned encoding cDNA. This may be performed by obtaining an amplified human α -N-acetylgalactosaminidase fragment capable of use as a screening probe, screening a chicken liver cDNA library, such as the one described hereinabove, using the amplified human
20 α -N-acetylgalactosaminidase fragment as a probe so as to obtain the cDNA sequence of the chicken liver cDNA library which encodes chicken liver α -N-acetylgalactosaminidase, sequencing the encoding DNA, cloning the encoding DNA and expressing chicken liver α -N-acetylgalactosaminidase enzyme
25 from the cloned encoding cDNA. Alternatively, screening can be

performed using antibodies which recognize chicken liver α -N-acetylgalactosaminidase.

Methods which are well known to those skilled in the art can be used to construct expression vectors containing the chicken liver α -N-acetylgalactosaminidase coding sequence, with appropriate transcriptional/translational signals for expression of the enzyme in the corresponding expression systems. Appropriate organisms, cell types and expression systems include: cell-free systems such as a rabbit reticulocyte lysate system, prokaryotic bacteria, such as E. coli, eukaryotic cells, such as yeast, insect cells, mammalian cells (including human hepatocytes or Chinese hamster ovary (CHO) cells), plant cells or systems, and animal systems including oocytes and transgenic animals.

The entire chicken liver α -N-acetylgalactosaminidase coding sequence or functional fragments of functional equivalents thereof may be used to construct the above expression vectors for production of functionally active enzyme in the corresponding expression system. Due to the degeneracy of the DNA code, it is anticipated that other DNA sequences which encode substantially the same amino acid sequence may be used. Additionally, changes to the DNA coding sequence which alter the amino acid sequence of the chicken liver α -N-acetylgalactosaminidase enzyme may be introduced which result in the expression of functionally active enzyme. In

particular, amino acid substitutions may be introduced which are based on similarity to the replaced amino acids, particularly with regard to the charge, polarity, hydrophobicity, hydrophilicity, and size of the side chains of the amino acids.

Once a recombinant chicken liver α -N-acetylgalactosaminidase enzyme is cloned and expressed, said enzyme can be used to remove A antigens from the surface of cells in blood products. Methods of utilizing chicken liver α -N-acetylgalactosaminidase to remove A antigens from the surface of erythrocytes can be found in U.S. Patent No. 4,609,627 issued September 2, 1986 to Goldstein, entitled "Enzymatic Conversion of Certain Sub-type A and AB Erythrocytes", which is incorporated herein by reference. Sub-type A antigens can be removed from the surface of erythrocytes by contacting the erythrocytes with the recombinant chicken liver α -N-acetylgalactosaminidase enzyme of this invention for a period of time sufficient to remove the A antigens from the surface of the erythrocytes.

20

EXAMPLE

Isolation and Characterization of the Chicken Liver cDNA Clone

Chicken liver α -N-acetylgalactosaminidase was purified to homogeneity. The enzyme was a glycoprotein with a molecular weight of 80 kDa, and was dissociated into two

25

identical subunits at pH 7.5. Its optimal pH for cleavage of the synthetic p-nitrophenyl- α -N-acetylgalactosaminyl-pyranoside substrate was 3.65 and the activity dropped sharply when the pH was raised above 7. The N-terminal sequence obtained from the purified chicken liver α -N-acetylgalactosaminidase showed a strong homology with the corresponding sequence deduced from the human α -N-acetylgalactosaminidase cDNA clone described in Tsuji et al., and Wang et al.

10 In order to isolate and characterize the cDNA clone for chicken liver α -N-acetylgalactosaminidase, two oligonucleotides, corresponding to nucleotides 688 to 705 and 1219 to 1236 of the human α -N-acetylgalactosaminidase sequence published by Wang, et al. were synthesized. Using
15 human placental mRNA (Clontech) as a template, the specific cDNA was made from the downstream (C-terminal) oligonucleotide. Next, a DNA fragment corresponding to human α -N-acetylgalactosaminidase residues from 688 to 1236 was amplified from the cDNA by the hot-start PCR technique. The
20 PCR reaction mixture was preheated at 95°C for 5 minutes and maintained at 80°C while Taq DNA polymerase (Promega) was added to reduce the possible non-specific annealing at lower temperature. 35 cycles of amplification was then carried out as follows: 94°C for 1 minute, 50°C for 2 minutes and 72°C for
25 3 minutes. The same conditions for PCR were applied in all of the following experiments.

The PCR-amplified fragment was then used as a radioactively-labeled probe in the screening of a chicken liver cDNA library (Stratagene) based on homology hybridization. The filters containing the library were hybridized with the probe
5 overnight at 42°C in a solution of 50% formamide, 5XSSPE, 5XDenhardt's, 0.1% SDS and 0.1 mg/ml salmon sperm DNA. The filters were then washed as follows:

1. 3 X SSC + 0.1% SDS, 20 min. room temperature
2. 2 X SSC + 0.1% SDS, 20 min. room temperature
- 10 3. 1 X SSC + 0.1% SDS, 20 min. 56°C
4. 1 X SSC + 0.1% SDS, 20 min. 56°C

The filters were autoradiographed overnight at -70°C. The positive clones were picked up for the second-round screening following the same procedure. In total, three
15 consecutive screenings were carried out in order to obtain a well-isolated positive clone.

From approximately one million plaques screened, one positive clone was successfully isolated. The sequencing data indicated that the clone consists of a 1.2 kb 3'-untranslated
20 region and a 0.7 kb coding region which is highly homologous to human α -N-acetylgalactosaminidase. In order to obtain the missing coding sequence, the library was rescreened by using the 1.9 kb cDNA clone as a probe. However, no positive clone was identified by this approach.

25

The upstream cDNA sequence was then obtained by applying multiple amplification (the nested PCR technique) of a second chicken liver cDNA library (Clontech). Figure 1 represents a diagram of the strategy used to clone and sequence the chicken liver α -N-acetylgalactosaminidase cDNA. The cDNA encoding chicken liver α -N-acetylgalactosaminidase contained a 1.2 kb coding region (slashed area) and a 1.2 kb 3' untranslated region. The arrows at the bottom of the diagram indicate the sequencing strategy. CL1, CL2 and CL3 are oligonucleotides used as primers for the nested PCR. CL1 and CL2 are located at position 924-941 nt and 736-753 nt, respectively (see Figure 2). According to the N-terminal sequence of native chicken liver enzyme, the oligonucleotide CL3 [5'-CTGGAGAAC(T)GGA(GC)CTGGCT(CA)CG] was designed taking into account chicken codon usage and "best guess".

In the first-round PCR amplification, the whole cDNA library was used as a template in the presence of one specific primer (CL1) (see Figure 1) and one universal primer derived from the library vector (5'-CTGGTAATGGTAGCGACC). A small aliquot from the above reaction was directly taken for the second-round amplification with a different set of primers. The primer CL2 had the sequence located upstream of CL1 (Figure 1) and the second primer, CL3, was designed based on the N-terminal amino acid sequence from purified chicken liver α -N-acetylgalactosaminidase (see Figure 1). A 750 bp

fragment was sequenced to eliminate any possible PCR artifacts. Since the 750 bp fragment overlapped with the 1.9 kb clone isolated by the library-screening, the two fragments were linked together by PCR to reconstitute the cDNA
5 encoding chicken liver α -N-acetylgalactosaminidase (Figure 1). The DNA sequencing was performed according to standard procedure, and the coding region was sequenced in both orientations.

10 The Cloned DNA Encodes Chicken
Liver α -N-Acetylgalactosaminidase

The authenticity of the cDNA clone was established by co-linearity of deduced amino acid sequences with N-terminal and CNBr-digested peptide sequences from purified chicken liver α -N-acetylgalactosaminidase. Figure 2 represents the nucleic
15 acid sequence and deduced amino acid sequence of the chicken liver α -N-acetylgalactosaminidase cDNA clone. The underlined regions in Figure 2 match sequences obtained from the N-terminus and CNBr-derived fragments of enzyme purified from chicken liver. The first 3 nucleotides, ATG, were added during
20 subcloning to serve as the translational initiation codon for protein expression. The polyadenylation signal (AATAAA) at positions 2299-2304 nt is double-underlined. The boxed sequence indicates potential sites for N-glycosylation. According to the cDNA, the mature protein of 405 amino acids
25 has a molecular mass of about 45 kDa, consistent with that of

the purified enzyme estimated by SDS-PAGE. Due to the cloning approach applied, the sequence at the 5' end of the cDNA corresponded to the N-terminal sequence of the mature enzyme isolated from chicken liver.

5 In order to express the chicken liver α -N-acetylgalactosaminidase in a rabbit reticulocyte lysate, the sequence from 1 to 1260 nucleotides which contained the coding region for chicken liver α -N-acetylgalactosaminidase was subcloned into the vector PCR-II (Invitrogen) in such an
10 orientation that the T7 promoter was located upstream of the insert. Since the N-terminus of the mature protein started with leucine, a translational initiation codon, ATG, was added during the subcloning construction. The construct was then used as a template in a transcription-translation coupled
15 system, TNT system (Promega), for protein expression according to the procedure recommended by the manufacturer.

 In order to produce the recombinant α -N-acetylgalactosaminidase in large quantities in bacteria and purify the enzyme in a single-step fashion, the cDNA was
20 subcloned into the EcoRI site of the pTrcHis vector (Invitrogen) for expression in E. coli. Because of the sequence in the vector, the expressed enzyme contained a polyhistidine-tag in its N-terminus, which permitted one step purification by affinity chromatography from crude cell lysates.

25

Figure 3 represents the expression of chicken liver α -N-acetylgalactosaminidase in bacteria and rabbit reticulocyte lysate as shown by Western blotting. Lane 1 through lane 4 demonstrate the results of expression in a rabbit reticulocyte lysate. The expression was carried out in lysate in the presence of ^{35}S -methionine with (lane 1) or without (lane 2) the expression plasmid. Next, 5 μl of the reaction sample was loaded to a 12% SDS-PAGE. The gel was dried and autoradiographed for 2 hours and a band of an apparent molecular weight of about 45KDa was visualized with the expression plasmid (lane 1, Figure 3). In order to confirm the authenticity of the expressed protein, a Western blot was performed using a polyclonal antibody raised against α -N-acetylgalactosaminidase purified from chicken liver. Using non-labelled methionine instead, the same expression reaction was performed for a Western blot (Promega) as shown in lanes 3 and 4, with and without the expression plasmid, respectively. As indicated in Figure 3, the antibody specifically recognized a band from the reaction with expression plasmid (lane 3), but not in the control (lane 4). Lane 5 shows the protein expressed in bacteria and recognized by the same antibody on Western blot. Lane 6 shows the α -N-acetylgalactosaminidase purified from chicken liver as a positive control. Molecular weight size marker (m) is indicated on the left. Hence, it was confirmed that the

isolated cDNA clone codes for the chicken liver α -N-acetylgalactosaminidase.

Comparison of the Cloned Chicken Liver
Sequence with other Enzyme Sequences

5 The chicken liver α -N-acetylgalactosaminidase sequence was compared with published sequences of other α -N-acetylgalactosaminidases and α -galactosidases which cleave α -galactose sugar groups. Figure 4 shows a homology comparison between various α -N-acetylgalactosaminidases and
10 α -galactosidases. Alignment was carried out using both the computer program PROSIS (Hitachi Software Engineering Corp., Ltd.) and manual arrangement. The amino acid sequences were deduced from cDNAs. Sequences I and II are of α -N-acetylgalactosaminidases from chicken liver and human
15 placenta, respectively. Sequences III, IV, V and VI represent α -galactosidase from human, yeast, Cyamopsis tetragonoloba and Aspergillus niger, respectively. Sequences IV and VI are truncated at the C-terminus, as indicated by **. Identical or conservatively substituted amino acid residues (five out of six
20 or more) among the aligned protein sequences are boxed. The numbers above the sequences indicate the relative position of each peptide sequence.

 The deduced amino acid sequence from chicken liver α -N-acetylgalactosaminidase cDNA shows approximately 80%
25 homology with the human α -N-acetylgalactosaminidase as

determined by PROSIS. This homology indicates the relatedness of the human and chicken liver enzymes, despite the differences in the specific characteristics of the enzymes, particularly with regard to cleavage of the Forsmann antigen, as has already
5 been described. Also, polyclonal antibodies raised against chicken liver α -N-acetylgalactosaminidase enzyme do not cross react with the human enzyme. The specific amino acids responsible for these differences remain to be elucidated.

Yamachi et al. (1990) reported that a human
10 α -N-acetylgalactosaminidase cDNA with an insertion of 70bp at the position corresponding to number 376 in Figure 4 was not enzymatically active in a transient expression study in COS cells. The data suggests that the open reading frame shift caused by this insertion in the C-terminal portion of the
15 molecule is responsible for the loss of enzymatic activity, indicating that amino acids in the C-terminal region may be essential for α -N-acetylgalactosaminidase enzyme activity.

By sequence similarity searching (BLAST) (Altschul et al. 1990) of available protein databases followed by sequence
20 alignment using the PROSIS computer program and manual arrangement, it was found that α -N-acetylgalactosaminidase is highly homologous to α -galactosidases from human, yeast, *Cyamopsis tetragonoloba* and *Aspergillus niger* (ranging from 55% to 68% at the amino acid level). The extent of the amino acid
25 sequence homology, as shown in Figure 4, suggests that these

two functionally specific glycosidases might have evolved from a common ancestral gene. Considering the high degree of similarities and the nature of their substrates it is possible that the two exoglycosidases share a similar catalytic mechanism and the critical amino acid residues involved in both active sites are well conserved. The addition of chicken liver α -N-acetylgalactosaminidase cDNA to the family provides further insight into regions of the molecule which are important for the substrate binding specificity and enzymatic activity. Given the availability of cloned enzymes from a number of sources, the active site and catalytic mechanisms of α -N-acetylgalactosaminidase and α -galactosidase enzymes may now be studied by means of cDNA deletion and site-directed mutagenesis.

15 Expression of Active Chicken Liver
 α -N-acetylgalactosaminidase in Yeast

 The first 48 nucleotides of human α -N-acetylgalactosaminidase cDNA (Wang, et al. 1990) which correspond to the signal peptide sequence, were linked to the cloned chicken liver α -N-acetylgalactosaminidase coding region by PCR. The PCR amplified product was subcloned directly into the vector PCR-II (Invitrogen). Two EcoRI sites flanking the insert were used to subclone the entire α -N-acetylgalactosaminidase cDNA into the yeast expression vector pYES2 (Invitrogen) in such an orientation that the GAL 1

promoter was located upstream of the insert. The GAL 1 promoter provides expression of the inserted cDNA clone under galactose inducing growth conditions in yeast.

The yeast vector constructs were transformed into the yeast strain, INVSCI (Invitrogen) using standard procedures. To confirm the expression of the chicken liver α -N-acetylgalactosaminidase in yeast, the total proteins from cell extract and culture supernatant were prepared and separated by 12% SDS-PAGE and a Western blot performed (by standard conditions) using the polyclonal antibody raised against purified chicken liver α -N-acetylgalactosaminidase. The transformed yeast cells were grown in medium without uracil (Bio 101, Inc.). After 0.2% galactose induction, the cells were centrifuged and protein extracts were prepared using glass bead disruption. The secreted proteins in the culture supernatant were concentrated with a Centricon-30 (Amicon Division, W.R. Grace & Co.). The Western blot results are depicted in Figure 5.

Lanes 1 and 8 of Figure 5 show the α -N-acetylgalactosaminidase purified from chicken liver. Lane 2 through lane 4 are cell extracts from the yeast transformed with three different pYES2 constructs: the vector alone (lane 2), chicken liver α -N-acetylgalactosaminidase cDNA coding region (lane 3), and the coding region plus signal sequence (lane 4). Lane 5 is the culture supernatant from

transformed yeast used in Lane 4. Lane 7 shows the molecular weight standard. As shown in Figure 5, while the protein without signal peptide was expressed within yeast cells (lane 3), the protein with a signal peptide sequence was predominantly secreted into the media (lane 5). The larger molecular weight of the secreted protein observed on the Western blot was presumably caused by overglycosylation, as was observed for the expression of guar α -galactosidase in yeast (Fellinger, et al. 1991).

To purify the expressed α -N-acetylgalactosaminidase, concentrated culture supernatant was applied to an affinity column containing aminocaproylgalactosylamine agarose. After washing the column, the bound fraction was eluted with buffer containing 50mM N-acetylgalactosamine. This eluate contains expressed α -N-acetylgalactosaminidase of similar molecular weight to that of the enzyme purified from chicken liver, as indicated in lane 6 in Figure 5.

The expressed enzyme eluted from the column demonstrates activity toward the synthetic substrate p-nitrophenyl- α -N-acetylgalactosaminyipyranoside at pH 3.6. Heavily glycosylated enzyme did not bind to the affinity column and showed no activity against synthetic substrate. All the data taken together demonstrate production, secretion and purification of enzymatically active chicken liver α -N-acetylgalactosaminidase in yeast cells.

Although the invention herein has been described with reference to particular embodiments, it is to be understood that these embodiments are merely illustrative of various aspects of the invention. Thus, it is to be understood that
5 numerous modifications may be made in the illustrative embodiments and other arrangements may be devised without departing from the spirit and scope of the invention.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Alex Zhu
Jack Goldstein
- 5 (ii) TITLE OF INVENTION: Recombinant α -N-Acetylgalactosaminidase Enzyme and cDNA Encoding Said Enzyme
- (iii) NUMBER OF SEQUENCES: 7
- 10 (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Amster, Rothstein & Ebenstein
(B) STREET: 90 Park Avenue
(C) CITY: New York
(D) STATE: New York
15 (E) COUNTRY: U.S.A.
(F) ZIP: 10016
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: 3.5 inch 1.44 Mb storage diskette
20 (B) COMPUTER: IBM PC Compatible
(C) OPERATING SYSTEM: MS-DOS
(D) SOFTWARE: Word Processor (ASCII)
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: Not yet assigned
(B) FILING DATE: Not yet assigned
25 (C) CLASSIFICATION: Not yet assigned
- (vii) PRIOR APPLICATION DATA: None
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Pasqualini, Patricia A.
(B) REGISTRATION NUMBER: 34,894
30 (C) REFERENCE/DOCKET NUMBER: 63475/12
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (212) 697-5995
(B) TELEFAX: (212) 286-0854 or 286-0082
(C) TELEX: TWX 710-581-4766

(2) INFORMATION FOR SEQ ID NO: 1:

- 5 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2319
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
(A) DESCRIPTION: cDNA to mRNA
- (iii) HYPOTHETICAL: no
- 10 (iv) ANTI-SENSE: yes
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
(A) ORGANISM: chicken liver
(B) STRAIN:
15 (C) INDIVIDUAL ISOLATE:
(D) DEVELOPMENTAL STAGE:
(E) HAPLOTYPE:
(F) TISSUE TYPE:
20 (G) CELL TYPE:
(H) CELL LINE:
(I) ORGANELLE:
- (vii) IMMEDIATE SOURCE: library
- (viii) POSITION IN GENOME: unknown
(A) CHROMOSOME SEGMENT:
25 (B) MAP POSITION:
(C) UNITS:
- (ix) FEATURE:
(A) NAME/KEY: chicken liver α -N-
acetylgalactosaminidase
(B) LOCATION:
30 (C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
(A) AUTHORS:
35 (B) TITLE:
(C) JOURNAL:
(D) VOLUME:
(F) PAGES:
(G) DATE:

(H) DOCUMENT NUMBER:
 (I) FILING DATE:
 (J) PUBLICATION DATE:
 (K) RELEVANT RESIDUES:

5	(xi)	SEQUENCE DESCRIPTION: SEQ ID NO: 1:	
		ATG CTG GAG AAC GGG CTG GCG CGG ACC CCG CCC ATG GGC TGG TTG GCC	48
		Met Leu Glu Asn Gly Leu Ala Arg Thr Pro Pro Met Gly Trp Leu Ala	
		TGG GAG CGG TTC CGC TGC AAC GTG AAC TGC CGG GAG GAC CCC CGC CAG	96
		Trp Glu Arg Phe Arg Cys Asn Val Asn Cys Arg Glu Asp Pro Arg Gln	
10		TGC ATC AGT GAG ATG CTC TTC ATG GAG ATG GCA GAC CGA ATA GCA GAG	144
		Cys Ile Ser Glu Met Leu Phe Met Glu Met Ala Asp Arg Ile Ala Glu	
		GAC GGC TGG AGG GAG CTG GGC TAC AAG TAC ATC AAT ATC GAT GAC TGC	192
		Asp Gly Trp Arg Glu Leu Gly Tyr Lys Tyr Ile Asn Ile Asp Asp Cys	
15		TGG GCC GCC AAG CAG CGT GAC ACT GAG GGG CGG CTG GTG CCT GAC CCC	240
		Trp Ala Ala Lys Gln Arg Asp Thr Glu Gly Arg Leu Val Pro Asp Pro	
		GAG AGG TTC CCC CGG GGC ATT AAG GCC TTG GCT GAC TAC GTT CAT GCC	288
		Glu Arg Phe Pro Arg Gly Ile Lys Ala Leu Ala Asp Tyr Val His Ala	
		CGA GGC TTG AAG CTG GGC ATT TAT GGC GAC CTG GGC AGA CTC ACC TGT	336
		Arg Gly Leu Lys Leu Gly Ile Tyr Gly Asp Leu Gly Arg Leu Thr Cys	
20		GGA GGC TAC CCA GGC ACC ACG CTG GAC CGT GTG GAG CAG GAC GCA CAG	384
		Gly Gly Tyr Pro Gly Thr Thr Leu Asp Arg Val Glu Gln Asp Ala Gln	
		ACC TTC GCT GAG TGG GGT GTG GAC ATG CTG AAG CTA GAT GGG TGC TAC	432
		Thr Phe Ala Glu Trp Gly Val Asp Met Leu Lys Leu Asp Gly Cys Tyr	
25		TCA TCG GGG AAG GAG CAG GCA CAG GGC TAC CCA CAA ATG GCA AGG GCC	480
		Ser Ser Gly Lys Glu Gln Ala Gln Gly Tyr Pro Gln Met Ala Arg Ala	
		TTG AAC GCC ACT GGC CGC CCC ATC GTC TAC TCC TGC AGC TGG CCA GCC	528
		Leu Asn Ala Thr Gly Arg Pro Ile Val Tyr Ser Cys Ser Trp Pro Ala	
		TAC CAG GGG GGG CTG CCT CCC AAG GTG AAC TAC ACT CTC CTG GGT GAG	576
		Tyr Gln Gly Gly Leu Pro Pro Lys Val Asn Tyr Thr Leu Leu Gly Glu	
30		ATC TGC AAC CTG TGG CGG AAC TAC GAT GAC ATC CAG GAC TCA TGG GAC	624
		Ile Cys Asn Leu Trp Arg Asn Tyr Asp Asp Ile Gln Asp Ser Trp Asp	
		AGC GTG CTT TCC ATC GTG GAC TGG TTC TTC ACA AAC CAG GAT GTG CTG	672
		Ser Val Leu Ser Ile Val Asp Trp Phe Phe Thr Asn Gln Asp Val Leu	
35		CAG CCG TTT GCT GGC CCT GGC CAC TGG AAT GAC CCA GAC ATG CTC ATC	720
		Gln Pro Phe Ala Gly Pro Gly His Trp Asn Asp Pro Asp Met Leu Ile	

	ATT	GGA	AAT	TTC	GGT	CTC	AGC	TAT	GAG	CAG	TCA	CGT	TCC	CAA	ATG	GCC	768
	Ile	Gly	Asn	Phe	Gly	Leu	Ser	Tyr	Glu	Gln	Ser	Arg	Ser	Gln	Met	Ala	
	TTG	TGG	ACC	ATT	ATG	GCA	GCT	CCA	CTC	CTC	ATG	TCC	ACC	GAC	CTG	CGC	816
	Leu	Trp	Thr	Ile	Met	Ala	Ala	Pro	Leu	Leu	Met	Ser	Thr	Asp	Leu	Arg	
5	ACT	ATC	TCG	CCG	AGT	GCC	AAG	AAG	ATT	CTG	CAG	AAC	CGC	CTG	ATG	ATC	864
	Thr	Ile	Ser	Pro	Ser	Ala	Lys	Lys	Ile	Leu	Gln	Asn	Arg	Leu	Met	Ile	
	CAG	ATA	AAC	CAG	GAC	CCC	TTG	GGA	ATC	CAG	GGG	CGC	AGG	ATC	ATC	AAG	912
	Gln	Ile	Asn	Gln	Asp	Pro	Leu	Gly	Ile	Gln	Gly	Arg	Arg	Ile	Ile	Lys	
10	GAG	GGA	TCC	CAC	ATT	GAG	GTG	TTC	CTG	CGC	CCG	CTG	TCA	CAG	GCT	GCC	960
	Glu	Gly	Ser	His	Ile	Glu	Val	Phe	Leu	Arg	Pro	Leu	Ser	Gln	Ala	Ala	
	AGT	GCC	CTG	GTC	TTC	TTC	AGC	CGG	AGG	ACA	GAC	ATG	CCC	TTC	CGC	TAC	1008
	Ser	Ala	Leu	Val	Phe	Phe	Ser	Arg	Arg	Thr	Asp	Met	Pro	Phe	Arg	Tyr	
	ACC	ACC	AGT	CTT	GCC	AAG	CTT	GGC	TTC	CCC	ATG	GGA	GCT	GCA	TAT	GAG	1056
	Thr	Thr	Ser	Leu	Ala	Lys	Leu	Gly	Phe	Pro	Met	Gly	Ala	Ala	Tyr	Glu	
15	GTG	CAA	GAC	GTG	TAC	AGT	GGG	AAG	ATC	ATC	AGT	GGC	CTG	AAG	ACA	GGA	1104
	Val	Gln	Asp	Val	Tyr	Ser	Gly	Lys	Ile	Ile	Ser	Gly	Leu	Lys	Thr	Gly	
	GAC	AAC	TTC	ACA	GTG	ATC	ATC	AAC	CCC	TCA	GGG	GTG	GTG	ATG	TGG	TAC	1152
	Asp	Asn	Phe	Thr	Val	Ile	Ile	Asn	Pro	Ser	Gly	Val	Val	Met	Trp	Tyr	
20	CTG	TGT	CCC	AAA	GCA	CTG	CTC	ATC	CAG	CAG	CAA	GCT	CCT	GGG	GGG	CCC	1200
	Leu	Cys	Pro	Lys	Ala	Leu	Leu	Ile	Gln	Gln	Gln	Ala	Pro	Gly	Gly	Pro	
	TCG	CGC	CTG	CCC	CTT	CTG	TGA	GGC	CCA	TGA	TTG	GGA	GCC	CTG	GGA	TAC	1248
	Ser	Arg	Leu	Pro	Leu	Leu	***										
	ATC	TCA	CCG	CTG	CTC	AAG	TGC	CTT	CTT	CTG	GTG	TGG	CTG	GGG	GAG	GAC	1296
	ATG	CAG	CTT	GCT	CCT	CTG	GCA	CCA	CCT	GAT	GAT	TTC	TAC	TCA	TTC	CAC	1344
25	GTG	AAG	CAG	GAC	TTC	TTG	TTA	CTC	CCT	CCT	GAG	AGC	ATG	CAA	AGC	GCT	1392
	CTG	AGG	TCC	TCC	TGT	GGA	AGA	GGA	GTG	TTC	CCA	GTG	ACC	ATC	CTT	TAG	1440
	GAC	CAG	ATG	TGG	TCA	CCT	TTT	TTC	CTT	TGC	TTG	GCT	TAG	GAC	AAA	GGG	1488
	CTG	TCC	ACA	GGC	TGC	ACC	CCT	CTT	CCC	AGG	CAC	CAT	CCC	CAG	ACC	AGG	1536
	AGC	TCC	TGG	GGC	CAG	GCT	GTC	TCT	GTC	TGG	CAG	CAG	GAT	CAG	CAG	GTA	1584
30	ACA	CCA	CTA	CAG	TGT	AGT	CCG	CAC	ATA	ATG	AAA	AAG	AAA	TCT	AAA	CAA	1632
	AAC	GTG	TGC	CAG	TAG	TGT	ACT	GAA	CCC	GCT	CTG	GTT	ACA	GCA	GAG	CAA	1680

```

AAC CTG AGT TGT CCA TGC ACA ATC CCA GTA TCC TCA CTG TGG TGT TAG 1728
CAT GAA AAA TTG CAG TCA CAG TGC ATT GTG CAC GAG TGG TGT CTG GAA 1776
GAT GCT GAT GCT TGT TCG TGG TGG TCT TAA GGT GGG AGA TGC TCA TGG 1824
GTG CTG GCC AAG TTG CAT CTC AAT CTT GTG AGG CTG AAC CTT CCA GCA 1872
5 TTT CTC AGG GAA AGG CTC TTC CTT TTA AAG GCA GCC TGC ACA AAT AGA 1920
AGG GGC TCA GAA GGA CGC ACG AGG AGG GGC TCA GGT GGG CCG TGC TCC 1968
CCT GAC CAC CCC AAG AGG GGT CAA CTA CTC ACC AAA ATC TAC CCC TTT 2016
CAA GGC CAG GTC AGC CCA GGG AGA CGC ACC CAA GGT TAA ACC TCA AAA 2064
CAG GAA ATC ACC CTA TTT TAA ATT AGT GAG AAA TTG AAC TTC CCC ATT 2112
10 CTA TTC AGA TGA GGG CTA GAA GCC CAC TCT CCT TAG AAG GCA CGT GGT 2160
GGA TTC CTG CCC CTT GCA GAG ACA TTG TGG TCT GAA GCA AGA TGC TGA 2208
ATG TGA TCT TTG CAG CGC TGG AAA TGA CAT GTC TGT TTC ATG CTT GTG 2256
TGG GAG ATG GCT TTG TTT TTG TGA TTT TGA CAA TTT AAC TGA AAT AAA 2304
AGG GAA GCA GAG GGG 2319

```

15 (3) INFORMATION FOR SEQ ID NO: 2:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 406
 (B) TYPE: amino acid
 (C) STRANDEDNESS: double
 20 (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
 (A) DESCRIPTION: cDNA to mRNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: yes
- 25 (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
 (A) ORGANISM: chicken liver
 (B) STRAIN:

5 (C) INDIVIDUAL ISOLATE:
 (D) DEVELOPMENTAL STAGE:
 (E) HAPLOTYPE:
 (F) TISSUE TYPE:
 (G) CELL TYPE:
 (H) CELL LINE:
 (I) ORGANELLE:

(vii) IMMEDIATE SOURCE: library

10 (viii) POSITION IN GENOME: unknown
 (A) CHROMOSOME SEGMENT:
 (B) MAP POSITION:
 (C) UNITS:

15 (ix) FEATURE:
 (A) NAME/KEY: chicken liver α -N-acetylgalactosaminidase
 (B) LOCATION:
 (C) IDENTIFICATION METHOD:
 (D) OTHER INFORMATION:

20 (x) PUBLICATION INFORMATION:
 (A) AUTHORS:
 (B) TITLE:
 (C) JOURNAL:
 (D) VOLUME:
 25 (F) PAGES:
 (G) DATE:
 (H) DOCUMENT NUMBER:
 (I) FILING DATE:
 (J) PUBLICATION DATE:
 (K) RELEVANT RESIDUES:

30 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Met	Leu	Glu	Asn	Gly	Leu	Ala	Arg	Thr	Pro	Pro	Met	Gly	Trp	Leu	Ala	16	
Trp	Glu	Arg	Phe	Arg	Cys	Asn	Val	Asn	Cys	Arg	Glu	Asp	Pro	Arg	Gln	32	
Cys	Ile	Ser	Glu	Met	Leu	Phe	Met	Glu	Met	Ala	Asp	Arg	Ile	Ala	Glu	48	
Asp	Gly	Trp	Arg	Glu	Leu	Gly	Tyr	Lys	Tyr	Ile	Asn	Ile	Asp	Asp	Cys	64	
35	Trp	Ala	Ala	Lys	Gln	Arg	Asp	Thr	Glu	Gly	Arg	Leu	Val	Pro	Asp	Pro	80
Glu	Arg	Phe	Pro	Arg	Gly	Ile	Lys	Ala	Leu	Ala	Asp	Tyr	Val	His	Ala	96	
Arg	Gly	Leu	Lys	Leu	Gly	Ile	Tyr	Gly	Asp	Leu	Gly	Arg	Leu	Thr	Cys	112	

Gly Gly Tyr Pro Gly Thr Thr Leu Asp Arg Val Glu Gln Asp Ala Gln 128
 Thr Phe Ala Glu Trp Gly Val Asp Met Leu Lys Leu Asp Gly Cys Tyr 144
 Ser Ser Gly Lys Glu Gln Ala Gln Gly Tyr Pro Gln Met Ala Arg Ala 160
 Leu Asn Ala Thr Gly Arg Pro Ile Val Tyr Ser Cys Ser Trp Pro Ala 176
 5 Tyr Gln Gly Gly Leu Pro Pro Lys Val Asn Tyr Thr Leu Leu Gly Glu 192
 Ile Cys Asn Leu Trp Arg Asn Tyr Asp Asp Ile Gln Asp Ser Trp Asp 208
 Ser Val Leu Ser Ile Val Asp Trp Phe Phe Thr Asn Gln Asp Val Leu 224
 Gln Pro Phe Ala Gly Pro Gly His Trp Asn Asp Pro Asp Met Leu Ile 240
 Ile Gly Asn Phe Gly Leu Ser Tyr Glu Gln Ser Arg Ser Gln Met Ala 256
 10 Leu Trp Thr Ile Met Ala Ala Pro Leu Leu Met Ser Thr Asp Leu Arg 272
 Thr Ile Ser Pro Ser Ala Lys Lys Ile Leu Gln Asn Arg Leu Met Ile 288
 Gln Ile Asn Gln Asp Pro Leu Gly Ile Gln Gly Arg Arg Ile Ile Lys 304
 Glu Gly Ser His Ile Glu Val Phe Leu Arg Pro Leu Ser Gln Ala Ala 320
 Ser Ala Leu Val Phe Phe Ser Arg Arg Thr Asp Met Pro Phe Arg Tyr 336
 15 Thr Thr Ser Leu Ala Lys Leu Gly Phe Pro Met Gly Ala Ala Tyr Glu 352
 Val Gln Asp Val Tyr Ser Gly Lys Ile Ile Ser Gly Leu Lys Thr Gly 368
 Asp Asn Phe Thr Ile Val Ile Asn Pro Ser Gly Val Val Met Trp Tyr 384
 Leu Cys Pro Lys Ala Leu Leu Ile Gln Gln Gln Ala Pro Gly Gly Pro 400
 Ser Arg Leu Pro Leu Leu 406

20 (4) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 411
 (B) TYPE: amino acid
 (C) STRANDEDNESS: double
 25 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE:

- (A) DESCRIPTION: cDNA to mRNA

(iii) HYPOTHETICAL: no

(iv) ANTI-SENSE: yes

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

5 (A) ORGANISM: human

(B) STRAIN:

(C) INDIVIDUAL ISOLATE:

(D) DEVELOPMENTAL STAGE:

10 (E) HAPLOTYPE:

(F) TISSUE TYPE:

(G) CELL TYPE:

(H) CELL LINE:

(I) ORGANELLE:

(vii) IMMEDIATE SOURCE: library

15 (viii) POSITION IN GENOME: unknown

(A) CHROMOSOME SEGMENT:

(B) MAP POSITION:

(C) UNITS:

(ix) FEATURE:

20 (A) NAME/KEY: human α -N-acetylgalactosaminidase

(B) LOCATION:

(C) IDENTIFICATION METHOD:

(D) OTHER INFORMATION:

25 (x) PUBLICATION INFORMATION:

(A) AUTHORS: Wang et al

(B) TITLE: Human α -N-Acetylgalactosaminidase
Molecular Cloning, Nucleotide Sequence,
and Expression of a Full-Length cDNA

30 (C) JOURNAL: Journal of Biological Chemistry

(D) VOLUME: 265

(F) PAGES: 21859-21866

(G) DATE: 1990

(H) DOCUMENT NUMBER:

35 (I) FILING DATE:

(J) PUBLICATION DATE:

(K) RELEVANT RESIDUES:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

Met Leu Leu Lys Thr Val Leu Leu Leu Gly His Val Ala Gln Val Leu 16

40 Met Leu Asp Asn Gly Leu Leu Gln Thr Pro Pro Met Gly Trp Leu Ala 32

Trp Glu Arg Phe Arg Cys Asn Ile Asn Cys Asp Glu Asp Pro Lys Asn 48
 Cys Ile Ser Glu Gln Leu Phe Met Glu Met Ala Asp Arg Met Ala Gln 64
 Asp Gly Trp Arg Asp Met Gly Tyr Thr Tyr Leu Asn Ile Asp Asp Cys 80
 Trp Ile Gly Gly Arg Asp Ala Ser Gly Arg Leu Met Pro Asp Pro Lys 96
 5 Arg Phe Pro His Gly Ile Pro Phe Leu Ala Asp Tyr Val His Ser Leu 112
 Gly Leu Lys Leu Gly Ile Tyr Ala Asp Met Gly Asn Phe Thr Cys Met 128
 Gly Tyr Pro Gly Thr Thr Leu Asp Lys Val Val Gln Asp Ala Gln Thr 144
 Phe Ala Glu Trp Lys Val Asp Met Leu Lys Leu Asp Gly Cys Phe Ser 160
 Thr Pro Glu Glu Arg Ala Gln Gly Tyr Pro Lys Met Ala Ala Ala Leu 176
 10 Asn Ala Thr Gly Arg Pro Ile Ala Phe Ser Cys Ser Trp Pro Ala Tyr 192
 Glu Gly Gly Leu Pro Pro Arg Val Asn Tyr Ser Leu Leu Ala Asp Ile 208
 Cys Asn Leu Trp Arg Asn Tyr Asp Asp Ile Gln Asp Ser Trp Trp Ser 224
 Val Leu Ser Ile Leu Asn Trp Phe Val Glu His Gln Asp Ile Leu Gln 240
 Pro Val Ala Gly Pro Gly His Trp Asn Asp Pro Asp Met Leu Leu Ile 256
 15 Gly Asn Phe Gly Leu Ser Leu Glu Gln Ser Arg Ala Gln Met Ala Leu 272
 Trp Thr Val Leu Ala Ala Pro Leu Leu Met Ser Thr Asp Leu Arg Thr 288
 Ile Ser Ala Gln Asn Met Asp Ile Leu Gln Asn Pro Leu Met Ile Lys 304
 Ile Asn Gln Asp Pro Leu Gly Ile Gln Gly Arg Arg Ile His Lys Glu 320
 Lys Ser Leu Ile Glu Val Tyr Met Arg Pro Leu Ser Asn Lys Ala Ser 336
 20 Ala Leu Val Phe Phe Ser Cys Arg Thr Asp Met Pro Tyr Arg Tyr His 352
 Ser Ser Leu Gly Gln Leu Asn Phe Thr Gly Ser Ile Val Tyr Glu Ala 368
 Gln Asp Val Tyr Ser Gly Asp Ile Ile Ser Gly Leu Arg Asp Glu Thr 384
 Asn Phe Thr Ile Val Ile Asn Pro Ser Gly Val Val Met Trp Tyr Leu 400
 Tyr Pro Ile Lys Asn Leu Glu Met Ser Gln Gln 411

(5) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 429
(B) TYPE: amino acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
(A) DESCRIPTION: cDNA to mRNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: yes
- (v) FRAGMENT TYPE: -
- (vi) ORIGINAL SOURCE:
(A) ORGANISM: human
(B) STRAIN:
(C) INDIVIDUAL ISOLATE:
(D) DEVELOPMENTAL STAGE:
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
(H) CELL LINE:
(I) ORGANELLE:
- (vii) IMMEDIATE SOURCE: library
- (viii) POSITION IN GENOME: unknown
(A) CHROMOSOME SEGMENT:
(B) MAP POSITION:
(C) UNITS:
- (ix) FEATURE:
(A) NAME/KEY: human α -galactosidase
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION:
- (x) PUBLICATION INFORMATION:
(A) AUTHORS: Calhoun et al
(B) TITLE: Fabry Disease: Isolation of a cDNA
Clone Encoding Human α -Galactosidase A
(C) JOURNAL: Proceedings of the National Academy
of Science USA
(D) VOLUME: 82
(F) PAGES: 7364-7368
(G) DATE: 1985

(H) DOCUMENT NUMBER:
 (I) FILING DATE:
 (J) PUBLICATION DATE:
 (K) RELEVANT RESIDUES:

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

	Met Gln Leu Arg Asn Pro Glu Leu His Leu Gly Cys Ala Leu Ala Leu	16
	Arg Phe Leu Ala Leu Val Ser Trp Asp Ile Pro Gly Ala Arg Ala Leu	32
	Asp Asn Gly Leu Ala Arg Thr Pro Thr Met Gly Trp Leu His Trp Glu	48
	Arg Phe Met Cys Asn Leu Asp Cys Gln Glu Glu Pro Asp Ser Cys Ile	64
10	Ser Glu Lys Leu Phe Met Glu Met Ala Glu Leu Met Val Ser Glu Gly	80
	Trp Lys Asp Ala Gly Tyr Glu Tyr Leu Cys Ile Asp Asp Cys Trp Met	96
	Ala Pro Gln Arg Asp Ser Glu Gly Arg Leu Gln Ala Asp Pro Gln Arg	112
	Phe Pro His Gly Ile Arg Gln Leu Ala Asn Tyr Val His Ser Lys Gly	128
	Leu Lys Leu Gly Ile Tyr Ala Asp Val Gly Asn Lys Thr Cys Ala Gly	144
15	Phe Pro Gly Ser Phe Gly Tyr Tyr Asp Ile Asp Ala Gln Thr Phe Ala	160
	Asp Trp Gly Val Asp Leu Leu Lys Phe Asp Gly Cys Tyr Cys Asp Ser	176
	Leu Glu Asn Leu Ala Asp Gly Tyr Lys His Met Ser Leu Ala Leu Asn	192
	Arg Thr Gly Arg Ser Ile Val Tyr Ser Cys Glu Trp Pro Leu Tyr Met	208
	Trp Pro Phe Gln Lys Pro Asn Tyr Thr Glu Ile Arg Gln Tyr Cys Asn	224
20	His Trp Arg Asn Phe Ala Asp Ile Asp Asp Ser Trp Lys Ser Ile Lys	240
	Ser Ile Leu Asp Trp Thr Ser Phe Asn Gln Glu Arg Ile Val Asp Val	256
	Ala Gly Pro Gly Gly Trp Asn Asp Pro Asp Met Leu Ile Val Gly Asn	272
	Phe Gly Leu Ser Trp Asn Gln Gln Val Thr Gln Met Ala Leu Trp Ala	288
	Ile Met Ala Ala Pro Leu Phe Met Ser Asn Asp Leu Arg His Ile Ser	304
25	Pro Gln Ala Lys Ala Leu Leu Gln Asp Lys Asp Ile Val Ala Ile Asn	320
	Gln Asp Pro Leu Gly Lys Gln Gly Tyr Gln Leu Arg Gln Gly Asp Asn	336
	Phe Glu Val Trp Glu Arg Pro Leu Ser Gly Leu Ala Trp Ala Val Ala	352

Met Ile Asn Arg Gln Glu Ile Gly Gly Pro Arg Ser Tyr Thr Ile Ala 368
Val Ala Ser Leu Gly Lys Gly Val Ala Cys Asn Pro Ala Cys Phe Ile 384
Thr Gln Leu Leu Pro Val Lys Arg Lys Leu Gly Phe Tyr Glu Trp Thr 400
Ser Arg Leu Arg Ser His Ile Asn Pro Thr Gly Thr Val Leu Leu Gln 416
5 Leu Glu Asn Thr Met Gln Met Ser Leu Lys Asp Leu Leu 429

(6) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
10 (A) LENGTH: 438
(B) TYPE: amino acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE:
(A) DESCRIPTION: cDNA to mRNA
- (iii) HYPOTHETICAL: no
- 15 (iv) ANTI-SENSE: yes
- (v) FRAGMENT TYPE:
- (vi) ORIGINAL SOURCE:
20 (A) ORGANISM: yeast Saccharomyces cerevisiae
(B) STRAIN:
(C) INDIVIDUAL ISOLATE:
(D) DEVELOPMENTAL STAGE:
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
25 (H) CELL LINE:
(I) ORGANELLE:
- (vii) IMMEDIATE SOURCE: library
- (viii) POSITION IN GENOME: unknown
30 (A) CHROMOSOME SEGMENT:
(B) MAP POSITION:
(C) UNITS:
- (ix) FEATURE:
(A) NAME/KEY: yeast α -galactosidase (MEL1)
(B) LOCATION:

(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

5 (A) AUTHORS: Liljestrom
(B) TITLE: The Nucleotide Sequence of the
Yeast MEL1 Gene
(C) JOURNAL: Nucleic Acids Research
(D) VOLUME: 13
10 (F) PAGES: 7257-7268
(G) DATE: 1985
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES:

15 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

Met Phe Ala Phe Tyr Phe Leu Thr Ala Cys Ile Ser Leu Lys Gly Val	16
Phe Gly Val Ser Pro Ser Tyr Asn Gly Leu Gly Leu Thr Pro Gln Met	32
Gly Trp Asp Asn Trp Asn Thr Phe Ala Cys Asp Val Ser Glu Gln Leu	48
Leu Leu Asp Thr Ala Asp Arg Ile Ser Asp Leu Gly Leu Lys Asp Met	64
20 Gly Tyr Lys Tyr Ile Ile Leu Asp Asp Cys Trp Ser Ser Gly Arg Asp	80
Ser Asp Gly Phe Leu Val Ala Asp Glu Gln Lys Phe Pro Asn Gly Met	96
Gly His Val Ala Asp His Leu His Asn Asn Ser Phe Leu Phe Gly Met	112
Tyr Ser Ser Ala Gly Glu Tyr Thr Cys Ala Gly Tyr Pro Gly Ser Leu	128
Gly Arg Glu Glu Glu Asp Ala Gln Phe Phe Ala Asn Asn Arg Val Asp	144
25 Tyr Leu Lys Tyr Asp Asn Cys Tyr Asn Lys Gly Gln Phe Gly Thr Pro	160
Glu Ile Ser Tyr His Arg Tyr Lys Ala Met Ser Asp Ala Leu Asn Lys	176
Thr Gly Arg Pro Ile Phe Tyr Ser Leu Cys Asn Trp Gly Gln Asp Leu	192
Thr Phe Tyr Trp Gly Ser Gly Ile Ala Asn Ser Trp Arg Met Ser Gly	208
Asp Val Thr Ala Glu Phe Thr Arg Pro Asp Ser Arg Cys Pro Cys Asp	224
30 Gly Asp Glu Tyr Asp Cys Lys Tyr Ala Gly Phe His Cys Ser Ile Met	240
Asn Ile Leu Asn Lys Ala Ala Pro Met Gly Gln Asn Ala Gly Val Gly	256

```

      Gly Trp Asn Asp Leu Asp Asn Leu Glu Val Gly Val Gly Asn Leu Thr      272
      Asp Asp Glu Glu Lys Ala His Phe Ser Met Trp Ala Met Val Lys Ser      288
      Pro Leu Ile Ile Gly Ala Asn Val Asn Asn Leu Lys Ala Ser Ser Tyr      304
      Ser Ile Tyr Ser Gln Ala Ser Ile Val Ala Ile Asn Gln Asp Ser Asn      320
5    Gly Ile Pro Ala Thr Arg Val Trp Arg Tyr Tyr Val Ser Asp Thr Asp      336
      Glu Tyr Gly Gln Gly Glu Ile Gln Met Trp Ser Gly Pro Leu Asp Asn      352
      Gly Asp Gln Val Val Ala Leu Leu Asn Gly Gly Ser Val Ser Arg Pro      368
      Met Asn Thr Thr Leu Glu Glu Ile Phe Phe Asp Ser Asn Leu Gly Ser      384
      Lys Lys Leu Thr Ser Thr Trp Asp Ile Tyr Asp Leu Trp Ala Asn Arg      400
10   Val Asp Asn Ser Thr Ala Ser Ala Ile Leu Gly Arg Asn Lys Thr Ala      416
      Thr Gly Ile Leu Tyr Asn Ala Thr Glu Gln Ser Tyr Lys Asp Gly Leu      432
      Ser Lys Asn Asp Thr Arg **                                          438

```

(7) INFORMATION FOR SEQ ID NO: 6:

```

15      (i)    SEQUENCE CHARACTERISTICS:
           (A)   LENGTH:    411
           (B)   TYPE:      amino acid
           (C)   STRANDEDNESS: double
           (D)   TOPOLOGY:  linear

20      (ii)   MOLECULE TYPE:
           (A)   DESCRIPTION: cDNA to mRNA

           (iii)  HYPOTHETICAL:    no

           (iv)   ANTI-SENSE:      yes

           (v)    FRAGMENT TYPE:

25      (vi)   ORIGINAL SOURCE:
           (A)   ORGANISM:    guar plant Cyamopsis tetragonoloba
           (B)   STRAIN:
           (C)   INDIVIDUAL ISOLATE:
           (D)   DEVELOPMENTAL STAGE:
           (E)   HAPLOTYPE:
30      (F)   TISSUE TYPE:

```

(G) CELL TYPE:
 (H) CELL LINE:
 (I) ORGANELLE:

(vii) IMMEDIATE SOURCE: library

5 (viii) POSITION IN GENOME: unknown
 (A) CHROMOSOME SEGMENT:
 (B) MAP POSITION:
 (C) UNITS:

10 (ix) FEATURE:
 (A) NAME/KEY: guar α -galactosidase
 (B) LOCATION:
 (C) IDENTIFICATION METHOD:
 (D) OTHER INFORMATION:

15 (x) PUBLICATION INFORMATION:
 (A) AUTHORS: Overbeeke et al
 (B) TITLE: Cloning and Nucleotide Sequence of
 the α -Galactosidase cDNA From
Cyamopsis tetragonoloba (guar)
 (C) JOURNAL: Plant Molecular Biology
 20 (D) VOLUME: 13
 (F) PAGES: 541-550
 (G) DATE: 1989
 (H) DOCUMENT NUMBER:
 (I) FILING DATE:
 25 (J) PUBLICATION DATE:
 (K) RELEVANT RESIDUES:

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

Met	Ala	Thr	His	Tyr	Ser	Ile	Ile	Gly	Gly	Met	Ile	Ile	Val	Val	Leu	16
Leu	Met	Ile	Ile	Gly	Ser	Glu	Gly	Gly	Arg	Leu	Leu	Glu	Lys	Lys	Asn	32
30	Arg	Thr	Ser	Ala	Glu	Ala	Glu	His	Tyr	Asn	Val	Arg	Arg	Tyr	Leu	48
Glu	Asn	Gly	Leu	Gly	Gln	Thr	Pro	Pro	Met	Gly	Trp	Asn	Ser	Trp	Asn	64
His	Phe	Gly	Cys	Asp	Ile	Asn	Glu	Asn	Val	Val	Arg	Glu	Thr	Ala	Asp	80
Ala	Met	Val	Ser	Thr	Gly	Leu	Ala	Ala	Leu	Gly	Tyr	Gln	Tyr	Ile	Asn	96
Leu	Asp	Asp	Cys	Trp	Ala	Glu	Leu	Asn	Arg	Asp	Ser	Glu	Gly	Asn	Met	112
35	Val	Pro	Asn	Ala	Ala	Ala	Phe	Pro	Ser	Gly	Ile	Lys	Ala	Leu	Ala	128
Tyr	Val	His	Ser	Lys	Gly	Leu	Lys	Leu	Gly	Val	Tyr	Ser	Asp	Ala	Gly	144

```

Asn Gln Thr Cys Ser Lys Arg Met Pro Gly Ser Leu Gly His Glu Glu 160
Gln Asp Ala Lys Thr Phe Ala Ser Trp Gly Val Asp Tyr Leu Lys Tyr 176
Asp Asn Cys Glu Asn Leu Gly Ile Ser Val Lys Glu Arg Tyr Pro Pro 192
Met Gly Lys Ala Leu Leu Ser Ser Gly Arg Pro Ile Phe Phe Ser Met 208
5 Cys Glu Trp Gly Trp Glu Asp Pro Gln Ile Trp Ala Lys Ser Ile Gly 224
Asn Ser Trp Arg Thr Thr Gly Asp Ile Glu Asp Asn Trp Asn Ser Met 240
Thr Ser Ile Ala Asp Ser Asn Asp Lys Trp Ala Ser Tyr Ala Gly Pro 256
Gly Gly Trp Asn Asp Pro Asp Met Leu Glu Val Gly Asn Gly Gly Met 272
Thr Thr Glu Glu Tyr Arg Ser His Phe Ser Ile Trp Ala Leu Ala Lys 288
10 Ala Pro Leu Leu Val Gly Cys Asp Ile Arg Ala Met Asp Asp Thr Thr 304
His Glu Leu Ile Ser Asn Ala Glu Ile Val Ala Val Asn Gln Asp Lys 320
Leu Gly Val Gln Gly Lys Lys Val Lys Ser Thr Asn Asp Leu Glu Val 336
Trp Ala Gly Pro Leu Ser Asp Asn Lys Val Ala Val Ile Leu Trp Asn 352
Arg Ser Ser Ser Arg Ala Thr Val Thr Ala Ser Trp Ser Asp Ile Gly 368
15 Leu Gln Gln Gly Thr Thr Val Asp Ala Arg Asp Leu Trp Glu His Ser 384
Thr Gln Ser Leu Val Ser Gly Glu Ile Ser Ala Glu Ile Asp Ser His 400
Ala Cys Lys Met Tyr Val Leu Thr Pro Arg Ser 411

```

(8) INFORMATION FOR SEQ ID NO: 7:

- 20 (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 447
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear
- 25 (ii) MOLECULE TYPE:
- (A) DESCRIPTION: cDNA to mRNA
- (iii) HYPOTHETICAL: no
- (iv) ANTI-SENSE: yes

(v) FRAGMENT TYPE:

(vi) ORIGINAL SOURCE:

5 (A) ORGANISM: Aspergillus niger
(B) STRAIN:
(C) INDIVIDUAL ISOLATE:
(D) DEVELOPMENTAL STAGE:
(E) HAPLOTYPE:
(F) TISSUE TYPE:
(G) CELL TYPE:
10 (H) CELL LINE:
(I) ORGANELLE:

(vii) IMMEDIATE SOURCE: library

(viii) POSITION IN GENOME: unknown

15 (A) CHROMOSOME SEGMENT:
(B) MAP POSITION:
(C) UNITS:

(ix) FEATURE:

20 (A) NAME/KEY: Aspergillus niger α -galactosidase
(B) LOCATION:
(C) IDENTIFICATION METHOD:
(D) OTHER INFORMATION:

(x) PUBLICATION INFORMATION:

25 (A) AUTHORS: den Herder et al
(B) TITLE: Cloning and Expression of a Member
of the Aspergillus niger Gene Family
Encoding α -Galactosidase
(C) JOURNAL: Molecular and General Genetics
(D) VOLUME: 233
(F) PAGES: 404-410
30 (G) DATE: 1992
(H) DOCUMENT NUMBER:
(I) FILING DATE:
(J) PUBLICATION DATE:
(K) RELEVANT RESIDUES:

35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

Met Ile Gln Gly Leu Glu Ser Ile Met Asn Gln Gly Thr Lys Arg Ile 16
Leu Leu Ala Ala Thr Leu Ala Ala Thr Pro Trp Gln Val Tyr Gly Ser 32
Ile Glu Gln Pro Ser Leu Leu Pro Thr Pro Pro Met Gly Pro Asn Asn 48
Trp Ala Arg Phe Met Cys Asp Leu Asn Glu Thr Leu Phe Thr Glu Thr 64
40 Ala Asp Thr Met Ala Ala Asn Gly Leu Arg Asp Ala Gly Tyr Asn Arg 80

	Ile Asn Leu Asp Asp Cys Trp Met Ala Tyr Gln Arg Ser Asp Asn Gly	96
	Ser Leu Gln Trp Asn Thr Thr Lys Phe Pro His Gly Leu Pro Trp Leu	112
	Ala Lys Tyr Val Lys Ala Lys Gly Phe His Phe Gly Ile Tyr Glu Asp	128
	Ser Gly Asn Met Thr Cys Gly Gly Tyr Pro Gly Ser Tyr Asn His Glu	144
5	Glu Gln Asp Ala Asn Thr Phe Ala Ser Trp Gly Ile Asp Tyr Leu Lys	160
	Leu Asp Gly Cys Asn Val Tyr Ala Thr Gln Gly Arg Thr Leu Glu Glu	176
	Glu Tyr Lys Gln Arg Tyr Gly His Trp His Gln Val Leu Ser Lys Met	192
	Gln His Pro Leu Ile Phe Ser Glu Ser Ala Pro Ala Tyr Phe Ala Gly	208
	Thr Asp Asn Asn Thr Asp Trp Tyr Thr Val Met Asp Trp Val Pro Ile	224
10	Tyr Gly Glu Leu Ala Arg His Ser Thr Asp Ile Leu Val Tyr Ser Gly	240
	Ala Gly Ser Ala Trp Asp Ser Ile Met Asn Asn Tyr Asn Tyr Asn Thr	256
	Leu Leu Ala Arg Tyr Gln Arg Pro Gly Tyr Phe Asn Asp Pro Asp Phe	272
	Leu Ile Pro Asp His Pro Gly Leu Thr Ala Asp Glu Lys Arg Ser His	288
	Phe Ala Leu Trp Ala Ser Phe Ser Ala Pro Leu Ile Ile Ser Ala Tyr	304
15	Ile Pro Ala Leu Ser Lys Asp Glu Ile Ala Phe Leu Ile Asn Glu Ala	320
	Leu Ile Ala Val Asn Gln Asp Pro Leu Ala Gln Gln Ala Thr Leu Ala	336
	Ser Arg Asp Asp Thr Leu Asp Ile Leu Thr Arg Ser Leu Ala Asn Gly	352
	Asp Arg Leu Leu Thr Val Leu Asn Lys Gly Asn Thr Thr Val Thr Arg	368
	Asp Ile Pro Val Gln Trp Leu Gly Leu Thr Glu Thr Asp Cys Thr Tyr	384
20	Thr Ala Glu Asp Leu Trp Asp Gly Lys Thr Gln Lys Ile Ser Asp His	400
	Ile Lys Ile Glu Leu Ala Ser His Ala Thr Ala Val Phe Arg Leu Ser	416
	Leu Pro Gln Gly Cys Ser Ser Val Val Pro Thr Gly Leu Val Phe Asn	432
	Thr Ala Ser Gly Asn Cys Leu Thr Ala Ala Ser Asn Ser Ser Val **	447

49

INDICATIONS RELATING TO A DEPOSITED MICROORGANISM

(PCT Rule 136a)

A. The indications made below relate to the microorganism referred to in the description on page <u>16</u> , line <u>1-5</u>	
B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet <input type="checkbox"/>	
Name of depository institution American Type Culture Collection	
Address of depository institution (including postal code and country) 12301 Parklawn Drive Rockville, Maryland 20852 United States of America	
Date of deposit 17 March 1993	Accession Number ATCC No. 75434
C. ADDITIONAL INDICATIONS (leave blank if not applicable) This information is contained on an additional sheet <input type="checkbox"/>	
A DNA vector containing a nucleic acid sequence which encodes chicken liver α -N-acetylgalactosaminidase enzyme.	
D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE (If the indications are not for all designated States)	
E. SEPARATE FURNISHING OF INDICATIONS (leave blank if not applicable)	
The indications listed below will be submitted to the International Bureau later (specify the general nature of the indications e.g., "Accession Number of Deposit")	
For receiving Office use only <input checked="" type="checkbox"/> This sheet was received with the international application Authorized officer <i>Ignacio L. Lirio</i>	For International Bureau use only <input type="checkbox"/> This sheet was received by the International Bureau via Authorized officer

Form PCT/IB/71.14 (July 1992)

WE CLAIM:

1. A recombinant chicken liver α -N-acetylgalactosaminidase enzyme.
2. A recombinant chicken liver α -N-acetylgalactosaminidase enzyme having a molecular weight of about 45 kDa.
3. A recombinant chicken liver α -N-acetylgalactosaminidase enzyme which is immunoreactive with an antibody specific for chicken liver α -N-acetylgalactosaminidase.
4. A recombinant chicken liver α -N-acetylgalactosaminidase enzyme having about 80% amino acid sequence homology with human α -N-acetylgalactosaminidase enzyme.
5. A recombinant chicken liver α -N-acetylgalactosaminidase enzyme including the amino acid sequence depicted in Figure 2, from amino acid number 1 to amino acid number 406.
6. A recombinant vector containing a nucleotide sequence encoding chicken liver α -N-acetylgalactosaminidase.
7. The recombinant vector of Claim 6 which includes the nucleotide sequence depicted in Figure 2, from nucleotide number 1 to nucleotide number 1218.
8. The DNA vector containing a sequence which encodes chicken liver α -N-acetylgalactosaminidase deposited

on March 17, 1993 with the American Type Culture Collection, Rockville, Maryland, and catalogued as ATCC No. 75434.

9. A method of producing chicken liver α -N-acetylgalactosaminidase comprising culturing a cell
5 transformed with a recombinant vector containing a nucleotide sequence encoding chicken liver α -N-acetylgalactosaminidase, and recovering chicken liver α -N-acetylgalactosaminidase from said culture.

10. The method of Claim 9 wherein the recombinant
10 vector includes the nucleotide sequence depicted in Figure 2, from nucleotide number 1 to nucleotide number 1218.

11. The method of Claim 9 which further comprises the step of purifying active chicken liver α -N-acetylgalactosaminidase enzyme from said culture by
15 affinity column.

12. The method of Claim 11 wherein the affinity column is aminocaproylgalactosylamine agarose.

13. A method of removing A antigens from the surface of erythrocytes comprising contacting said erythrocytes with
20 recombinant chicken liver α -N-acetylgalactosaminidase enzyme for a period of time sufficient to remove said A antigens from the surface of said erythrocytes.

14. The method of Claim 13 wherein the recombinant chicken liver α -N-acetylgalactosaminidase enzyme has a
25 molecular weight of about 45 kDa.

15. The method of Claim 13 wherein the recombinant chicken liver α -N-acetylgalactosaminidase enzyme is immunoreactive with an antibody specific for chicken liver α -N-acetylgalactosaminidase.

5 16. The method of Claim 13 wherein the recombinant chicken liver α -N-acetylgalactosaminidase enzyme has about 80% amino acid sequence homology with human α -N-acetylgalactosaminidase.

10 17. The method of Claim 13 wherein the recombinant chicken liver α -N-acetylgalactosaminidase enzyme includes the amino acid sequence depicted in Figure 2, from amino acid number 1 to amino acid number 406.

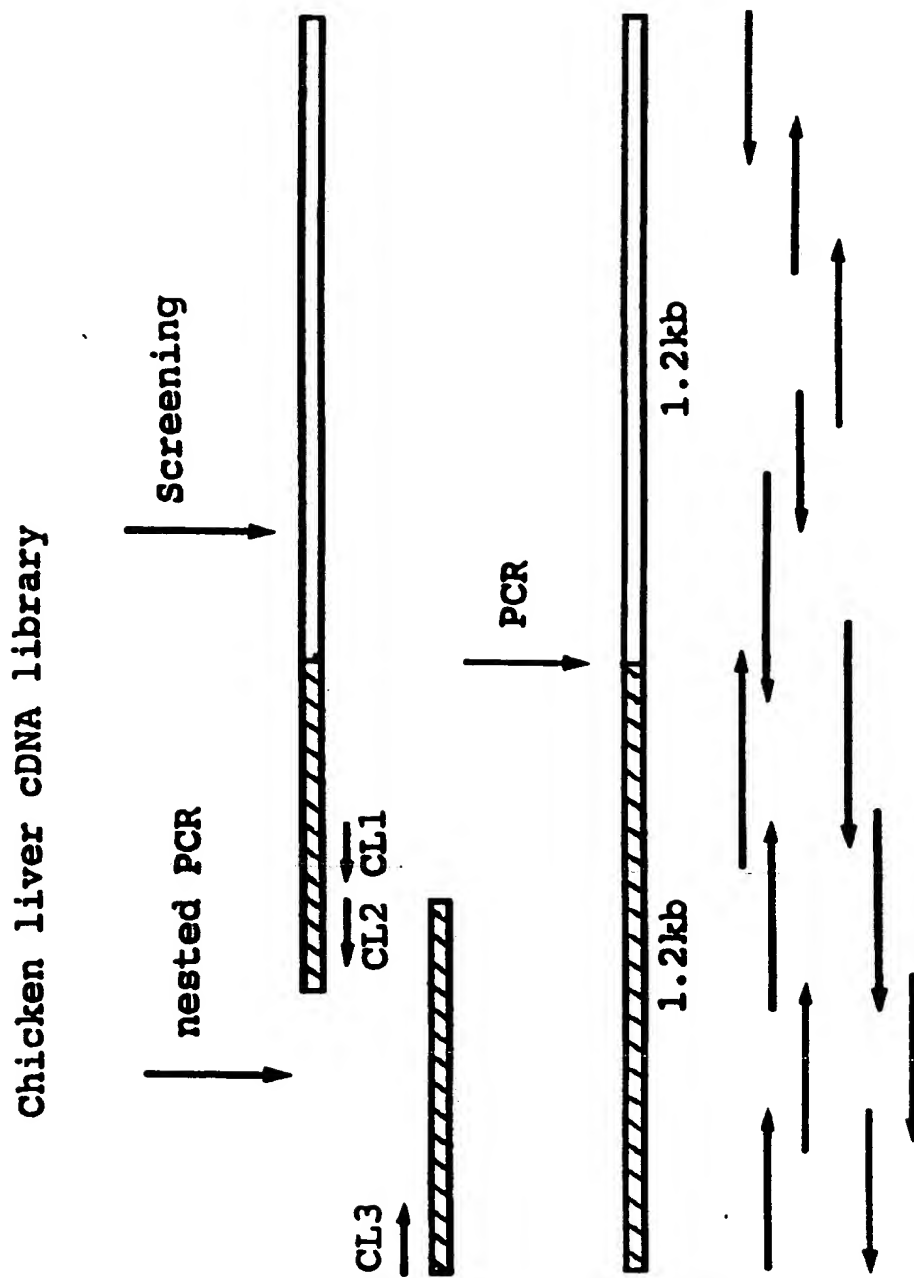


FIG. 1

- 2 / 6 -

ATG	CTG	GAG	AAC	GGG	CTG	GCG	CGG	ACC	CCG	CCC	ATG	GGC	TGG	TTG	GCC	48
Met	Leu	Glu	Asn	Gly	Leu	Ala	Arg	Thr	Pro	Pro	Met	Gly	Trp	Leu	Ala	
TGG	GAG	CGG	TTC	CGC	TGC	AAC	GTG	AAC	TGC	CGG	GAG	GAC	CCC	CGC	CAG	96
Trp	Glu	Arg	Phe	Arg	Cys	Asn	Val	Asn	Cys	Arg	Glu	Asp	Pro	Arg	Gln	
TGC	ATC	AGT	GAG	ATG	CTC	TTC	ATG	GAG	ATG	GCA	GAC	CGA	ATA	GCA	GAG	144
Cys	Ile	Ser	Glu	Met	Leu	Phe	Met	Glu	Met	Ala	Asp	Arg	Ile	Ala	Glu	
GAC	GGC	TGG	AGG	GAG	CTG	GGC	TAC	AAG	TAC	ATC	AAT	ATC	GAT	GAC	TGC	192
Asp	Gly	Trp	Arg	Glu	Leu	Gly	Tyr	Lys	Tyr	Ile	Asn	Ile	Asp	Asp	Cys	
TGG	GCC	GCC	AAG	CAG	CGT	GAC	ACT	GAG	GGG	CGG	CTG	GTG	CCT	GAC	CCC	240
Trp	Ala	Ala	Lys	Gln	Arg	Asp	Thr	Glu	Gly	Arg	Leu	Val	Pro	Asp	Pro	
GAG	AGG	TTC	CCC	CGG	GGC	ATT	AAG	GCC	TTG	GCT	GAC	TAC	GTT	CAT	GCC	288
Glu	Arg	Phe	Pro	Arg	Gly	Ile	Lys	Ala	Leu	Ala	Asp	Tyr	Val	His	Ala	
CGA	GGC	TTG	AAG	CTG	GGC	ATT	TAT	GGC	GAC	CTG	GGC	AGA	CTC	ACC	TGT	336
Arg	Gly	Leu	Lys	Leu	Gly	Ile	Tyr	Gly	Asp	Leu	Gly	Arg	Leu	Thr	Cys	
GGA	GGC	TAC	CCA	GGC	ACC	ACG	CTG	GAC	CGT	GTG	GAG	CAG	GAC	GCA	CAG	384
Gly	Gly	Tyr	Pro	Gly	Thr	Thr	Leu	Asp	Arg	Val	Glu	Gln	Asp	Ala	Gln	
ACC	TTC	GCT	GAG	TGG	GGT	GTG	GAC	ATG	CTG	AAG	CTA	GAT	GGG	TGC	TAC	432
Thr	Phe	Ala	Glu	Trp	Gly	Val	Asp	Met	Leu	Lys	Leu	Asp	Gly	Cys	Tyr	
TCA	TCG	GGG	AAG	GAG	CAG	GCA	CAG	GGC	TAC	CCA	CAA	ATG	GCA	AGG	GCC	480
Ser	Ser	Gly	Lys	Glu	Gln	Ala	Gln	Gly	Tyr	Pro	Gln	Met	Ala	Arg	Ala	
TTG	AAC	GCC	ACT	GGC	CGC	CCC	ATC	GTG	TAC	TCC	TGC	AGC	TGG	CCA	GCC	528
Leu	Asn	Ala	Thr	Gly	Arg	Pro	Ile	Val	Tyr	Ser	Cys	Ser	Trp	Pro	Ala	
TAC	CAG	GGG	GGG	CTG	CCT	CCC	AAG	GTG	AAC	TAC	ACT	CTC	CTG	GGT	GAG	576
Tyr	Gln	Gly	Gly	Leu	Pro	Pro	Lys	Val	Asn	Tyr	Thr	Leu	Leu	Gly	Glu	
ATC	TGC	AAC	CTG	TGG	CGG	AAC	TAC	GAT	GAC	ATC	CAG	GAC	TCA	TGG	GAC	624
Ile	Cys	Asn	Leu	Trp	Arg	Asn	Tyr	Asp	Asp	Ile	Gln	Asp	Ser	Trp	Asp	
AGC	GTG	CTT	TCC	ATC	GTG	GAC	TGG	TTC	TTC	ACA	AAC	CAG	GAT	GTG	CTG	672
Ser	Val	Leu	Ser	Ile	Val	Asp	Trp	Phe	Phe	Thr	Asn	Gln	Asp	Val	Leu	
CAG	CCG	TTT	GCT	GGC	CCT	GGC	CAC	TGG	AAT	GAC	CCA	GAC	ATG	CTC	ATC	720
Gln	Pro	Phe	Ala	Gly	Pro	Gly	His	Trp	Asn	Asp	Pro	Asp	Met	Leu	Ile	
ATT	GGA	AAT	TTC	GGT	CTC	AGC	TAT	GAG	CAG	TCA	CGT	TCC	CAA	ATG	GCC	768
Ile	Gly	Asn	Phe	Gly	Leu	Ser	Tyr	Glu	Gln	Ser	Arg	Ser	Gln	Met	Ala	
TTG	TGG	ACC	ATT	ATG	GCA	GCT	CCA	CTC	CTC	ATG	TCC	ACC	GAC	CTG	CGC	816
Leu	Trp	Thr	Ile	Met	Ala	Ala	Pro	Leu	Leu	Met	Ser	Thr	Asp	Leu	Arg	
ACT	ATC	TCG	CCG	AGT	GCC	AAG	AAG	ATT	CTG	CAG	AAC	CGC	CTG	ATG	ATC	864
Thr	Ile	Ser	Pro	Ser	Ala	Lys	Lys	Ile	Leu	Gln	Asn	Arg	Leu	Met	Ile	
CAG	ATA	AAC	CAG	GAC	CCC	TTG	GGA	ATC	CAG	GGG	CGC	AGG	ATC	ATC	AAG	912
Gln	Ile	Asn	Gln	Asp	Pro	Leu	Gly	Ile	Gln	Gly	Arg	Arg	Ile	Ile	Lys	
GAG	GGA	TCC	CAC	ATT	GAG	GTG	TTC	CTG	CGC	CCG	CTG	TCA	CAG	GCT	GCC	960
Glu	Gly	Ser	His	Ile	Glu	Val	Phe	Leu	Arg	Pro	Leu	Ser	Gln	Ala	Ala	
AGT	GCC	CTG	GTC	TTC	TTC	AGC	CGG	AGG	ACA	GAC	ATG	CCC	TTC	CGC	TAC	1008
Ser	Ala	Leu	Val	Phe	Phe	Ser	Arg	Arg	Thr	Asp	Met	Pro	Phe	Arg	Tyr	
ACC	ACC	AGT	CTT	GCC	AAG	CTT	GGC	TTC	CCC	ATG	GGA	GCT	GCA	TAT	GAG	1056
Thr	Thr	Ser	Leu	Ala	Lys	Leu	Gly	Phe	Pro	Met	Gly	Ala	Ala	Tyr	Glu	

FIG. 2A

SUBSTITUTE SHEET (RULE 26)

- 3 / 6 -

GTG	CAA	GAC	GTG	TAC	AGT	GGG	AAG	ATC	ATC	AGT	GGC	CTG	AAG	ACA	GGA	1104
Val	Gln	Asp	Val	Tyr	Ser	Gly	Lys	Ile	Ile	Ser	Gly	Leu	Lys	Thr	Gly	
GAC	AAC	TTC	ACA	GTG	ATC	ATC	AAC	CCC	TCA	GGG	GTG	GTG	ATG	TGG	TAC	1152
Asp	Asn	Phe	Thr	Val	Ile	Ile	Asn	Pro	Ser	Gly	Val	Val	Met	Trp	Tyr	
CTG	TGT	CCC	AAA	GCA	CTG	CTC	ATC	CAG	CAG	CAA	GCT	CCT	GGG	GGG	CCC	1200
Leu	Cys	Pro	Lys	Ala	Leu	Leu	Ile	Gln	Gln	Gln	Ala	Pro	Gly	Gly	Pro	
TCG	CGC	CTG	CCC	CTT	CTG	TGA	GGC	CCA	TGA	TTG	GGA	GCC	CTG	GGA	TAC	1248
Ser	Arg	Leu	Pro	Leu	Leu	***										
ATC	TCA	CCG	CTG	CTC	AAG	TGC	CTT	CTT	CTG	GTG	TGG	CTG	GGG	GAG	GAC	1296
ATG	CAG	CTT	GCT	CCT	CTG	GCA	CCA	CCT	GAT	GAT	TTC	TAC	TCA	TTC	CAC	1344
GTG	AAG	CAG	GAC	TTC	TTG	TTA	CTC	CCT	CCT	GAG	AGC	ATG	CAA	AGC	GCT	1392
CTG	AGG	TCC	TCC	TGT	GGA	AGA	GGA	GTG	TTC	CCA	GTG	ACC	ATC	CTT	TAG	1440
GAC	CAG	ATG	TGG	TCA	CCT	TTT	TTC	CTT	TGC	TTG	GCT	TAG	GAC	AAA	GGG	1488
CTG	TCC	ACA	GGC	TGC	ACC	CCT	CTT	CCC	AGG	CAC	CAT	CCC	CAG	ACC	AGG	1536
AGC	TCC	TGG	GGC	CAG	GCT	GTC	TCT	GTC	TGG	CAG	CAG	GAT	CAG	CAG	GTA	1584
ACA	CCA	CTA	CAG	TGT	AGT	CCG	CAC	ATA	ATG	AAA	AAG	AAA	TCT	AAA	CAA	1632
AAC	GTG	TGC	CAG	TAG	TGT	ACT	GAA	CCC	GCT	CTG	GTT	ACA	GCA	GAG	CAA	1680
AAC	CTG	AGT	TGT	CCA	TGC	ACA	ATC	CCA	GTA	TCC	TCA	CTG	TGG	TGT	TAG	1728
CAT	GAA	AAA	TTG	CAG	TCA	CAG	TGC	ATT	GTG	CAC	GAG	TGG	TGT	CTG	GAA	1776
GAT	GCT	GAT	GCT	TGT	TCG	TGG	TGG	TCT	TAA	GGT	GGG	AGA	TGC	TCA	TGG	1824
GTG	CTG	GCC	AAG	TTG	CAT	CTC	AAT	CTT	GTG	AGG	CTG	AAC	CTT	CCA	GCA	1872
TTT	CTC	AGG	GAA	AGG	CTC	TTC	CTT	TTA	AAG	GCA	GCC	TGC	ACA	AAT	AGA	1920
AGG	GGC	TCA	GAA	GGA	CGC	ACG	AGG	AGG	GGC	TCA	GGT	GGG	CCG	TGC	TCC	1968
CCT	GAC	CAC	CCC	AAG	AGG	GGT	CAA	CTA	CTC	ACC	AAA	ATC	TAC	CCC	TTT	2016
CAA	GGC	CAG	GTC	AGC	CCA	GGG	AGA	CGC	ACC	CAA	GGT	TAA	ACC	TCA	AAA	2064
CAG	GAA	ATC	ACC	CTA	TTT	TAA	ATT	AGT	GAG	AAA	TTG	AAC	TTC	CCC	ATT	2112
CTA	TTC	AGA	TGA	GGG	CTA	GAA	GCC	CAC	TCT	CCT	TAG	AAG	GCA	CGT	GGT	2160
GGA	TTC	CTG	CCC	CTT	GCA	GAG	ACA	TTG	TGG	TCT	GAA	GCA	AGA	TGC	TGA	2208
ATG	TGA	TCT	TTG	CAG	CGC	TGG	AAA	TGA	CAT	GTC	TGT	TTC	ATG	CTT	GTG	2256
TGG	GAG	ATG	GCT	TTG	TTT	TTG	TGA	TTT	TGA	CAA	TTT	AAC	TGA	AAT	AAA	2304
AGG	GAA	GCA	GAG	GGG												2319

FIG. 2B

SUBSTITUTE SHEET (RULE 26)

- 4 / 6 -

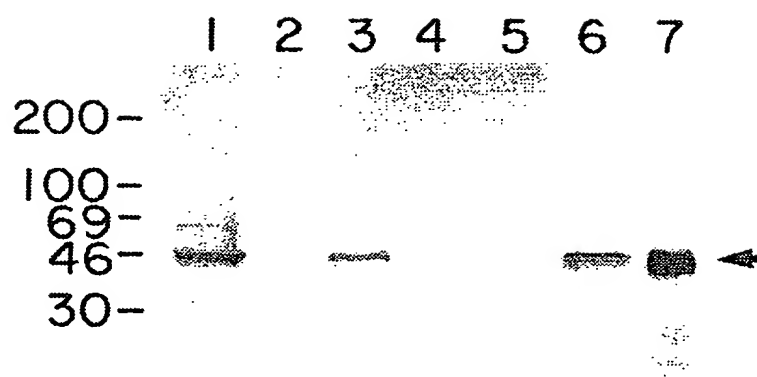


FIG. 3

- 5 / 6 -

	20	40	60
I			MLENGLARTPPMGWLAWERFR CNVN
II		MLLKTVLLLGHVAVQVIMLDNGILQTPPMGWLAWERFR CNIN	
III	MQLRNP EHLG CALALRFLALVSWDIPGARALDNGIARTPTMGWLHWERFMCNLD		
IV		MFAFYFLTACISLKG VFGVSPSYNGLGLTTPMGWDNWNIFACDV	
V	MATHYSIIGGMIIVVLLMIIGSEGGRLLEKKNR TSAEAEHYNVRRYLAENGLGQTTPPMGWNSWNHFGCDIN		
VI	MIQGLESIMNQGTKRILLAATLAATPWQVYGSIEQPSLLPTPPMGFNNAWRFMCDLN		

	80	100	120	140
I	CREDPRQCISEMLFMEMADRIAEDGWRELGYKYINIDDCWAAKQRDTEGRIMVDPERFFRGIKALADYVHAR			
II	CDEDPKNCISEQLFMEMADRIAQDGRDMGYTYINIDDCWIGG RDA SGRLMPDPKRFHFGIPFLADYVHSL			
III	CQEEPDCSISEKLFMEMAEIMVSEGWKDAGYEYICIDDCWMAPO RDSEGRLOADPQRFHFGIRQLANYVHSH			
IV	SEOLLIDTADRISDGLKIDMGYKYIILDDCWSSG RDSDFLVADEQKFFNGMGHVADELHNN			
V	ENVVREITADAMVSTGLAALGYQYINLDDCWAEINRDSEGNMVPNAAFHSGIKALADYVHSH			
VI	EILFTETADITMAANGLRDAGYNRIINLDDCWMAQORSINGSLQWNITKFFHGLPWLA KYVVKAK			

	160	180	200
I	GLKLG IYGDIGRLTCG GYPGTTLDRVEQDAQTFAEWGVDM LKLDGCYSSGKEQ		AQGYPQMARAL
II	GLKLG IYADMGNFICM GYPGTTLDKVVQDAQTFAEWKVDMLKLDGCFSTPEER		AQGYPKMAAAL
III	GLKLG IYADVGNKTCG GYPGSG FGYDIDDAQTFADWGVLLKFDGCYCD SLEN		LADGKXKMSIAL
IV	SFLFGMYSSAGEYTCA GYPGSG LGREEEDAQFFANNRVDYLKYDNCYNKGQFG		TPEISYHYKAMS DAL
V	GLKLG VYS DAGNOTCSKRM PGSG LGHEEQDAKTFASWGV DYLKYDNCENLGISV		KERYPPMGKAL
VI	GFHFGIYEDSGNMTCC GYPGSG YNHEEQDANTFASW GIDYLKLDGCNVYATQGR TLEEYKQRYGHWHQVL		

	220	240	260	280
I	NATGRPIVYS CSWPAYQGG LPPKVNYTLLGEICNLWRNYDDIQDSW SVLSIVDWFFTNQDVLQPF			
II	NATGRPIAFS CSWPAYEGGLPPRVNYSLLADICNLWRNYDDIQDSW SVLSILNWFVEHQDILQPV			
III	NRTGRSIVYS CEWPLYMWPFQ KPNYTEIRQYCNHWRNFADIDDSWKS IKSILDWTSFNQERIVDV			
IV	NKTGRPIFYSLONWGQDLTFYWGSGIA	NSWRMSGDVTAEFTRPDSRCPCDGD EYDCKYAGFHCSI		
V	LSSGRPIFFSMCEWGWEDPQIWAKSIG	NSWRITGDIEDNNSMTSIADSN DKWASY		
VI	SKMCHPLIFSESAPAYFAGTDNNTD WYTVMDWVPIY GELARHSTDILVYSGAGSAWDSIMNNY			

	300	320	340
I	AGPGHWNDPDM LIGNFGLSYEQSR SQMALWTIMAAPLLMSTD LRITISPAK KILONRL		
II	AGPGHWNDPDM LIGNFGLSYEQSR AQMALWTVLAAPLLMSTD LRITISAQNM DILO NPL		
III	AGPGHWNDPDM LIGNFGLSWNQOVITOMALWAIMAAPLLMSTD LRITISPOAKALLODKD		
IV	MNILNKAAPMGQNA GVGWNDLIDNLEVGVGNLTDEEKAHFSMWAMVKSPLIIGANVN NIKASSYSIYSCAS		
V	AGPGHWNDPDM LLEVGNNGMTIEEYRSHFSI WALAKAPLLVGCDIRAMDDTTHELISNAE		
VI	NYNTLLARYQRP GYFNDPDLIPDHFGLTAD EKRSHFALWASFSAPLIISAVIPALSKDEIAFLINEA		

	360	380	400	420
I	MIQINQDPLGIQGRRII KEKSHIEVFLRPLSQAASALVFFSRRT DMPFRYT TSLAKLGFP MGAAYEV Q			
II	MIKINQDPLGIQGRRII KEKSLIEVYMRPLSNKASALVFFSCRT DMPYRYHSSLGQLNFTG SIVYEA Q			
III	IVAINQDPLGKQGYQL RQGDNF EVWERPLSGLAWAVAMINRQEIGGPRSYTIAVASLGKG VACNPACFITQ			
IV	IVAINQDSNGIPATRVWRYVSDTDEYGGQEI QMWSGPLDNGDQVALLNGGSVSRPMNTTLEEIFFDSNLG			
V	IVAVNQDKLGVOGKKV KSTNDLEVWAGPLSDNKVAVILWNRSSSRATVTASWSDIGLQO GTTVDARDLWEH			
VI	LIAVNQDPLAQATLAS RQDTLDILTRSLANGDRLLTVLNKGNTTVTRDIPVQWLG LTETDCTYTAEDLWDG			

	440	460	480
I	DVYSGKIIISGLKTGDNFTIVINPSGVVMWYLCPKALLIQQQAPGGPSRLPLL		
II	DVYSGDIISGLRDETNTFTIVINPSGVVMWYLYPIKNLEMSQQ		
III	LLPVKRKLGFYEWT SRLRSHINPTGTVLLQLENTMQMSLKDLL		
IV	SKKLTSTWDIYDLWANRV DNSTASAILGRNKTATGILYNATEQSYKDGLSKNDTR*		
V	STQSLVSGEISAEIDSHACKMYVLTPRS		
VI	KTQKISDHIKIELASHATAVFR LSLPQGCSSVVTGLVFNTASGNCLTAASNSSV**		

FIG. 4

SUBSTITUTE SHEET (RULE 26)

- 6 / 6 -

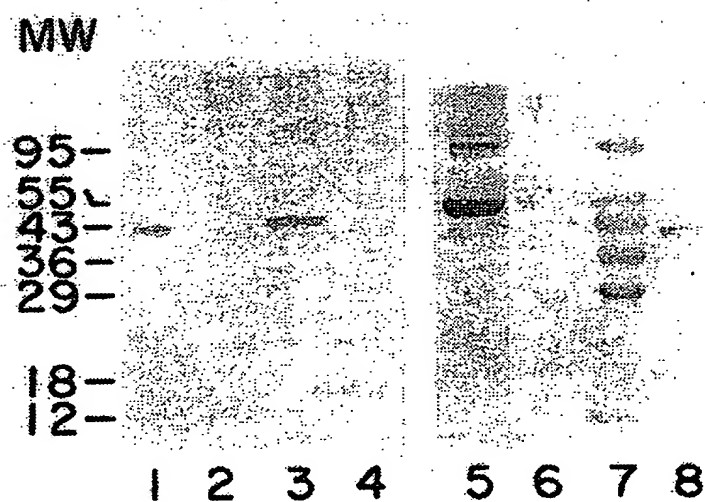


FIG. 5

SUBSTITUTE SHEET (RULE 26)

INTERNATIONAL SEARCH REPORT

International application N .
PCT/US94/03338

A. CLASSIFICATION OF-SUBJECT MATTER

IPC(5) :C12S 3/22; C12N 9/40, 15/10, 15/56, 15/63; A01N 1/02

US CL :435/208, 320.1, 69.1, 269, 2, 172.3; 536/23.2

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 435/208, 320.1, 69.1, 269, 2, 172.3; 536/23.2

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

APS, CA, MEDLINE, BIOSIS, EMBASE, LIFESCI, BIOTECHDS, WPIDS
search terms: acetylgalactosaminidase#, chicken liver, gene# or sequence#

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, X ----- Y	Biochimica et Biophysica Acta, Volume 1216, No. 2, issued November 1993, M.O. Davis et al., "Cloning and Sequence of a Chicken α -N-acetylgalactosaminidase Gene", pages 296-298, see entire document.	6-8 ----- 1-7, 9-17
P, X ----- Y	Gene, Volume 137, No. 2, issued December 1993, A. Zhu et al., "Cloning and Characterization of a cDNA Encoding Chicken Liver α -N-acetylgalactosaminidase, pages 309-314, see entire document.	1-12 ----- 13-17
P, X ----- Y	FASEB Journal, Volume 7, No. 7, issued 20 April 1993, A. Zhu et al., "Molecular Cloning and Characterization of a Chicken Liver α -N-Acetylgalactosaminidase", page A1251, see entire document.	1-12 ----- 13-17



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:	T	later documents published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
A document defining the general state of the art which is not considered to be of particular relevance	X	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
E earlier document published on or after the international filing date	Y	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
L document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	A	document member of the same patent family
O documents referring to an oral disclosure, use, exhibition or other means		
P documents published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

10 JUNE 1994

Date of mailing of the international search report

JUN 29 1994

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer

REBECCA PROUTY

Telephone N. (703)308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

INTERNATIONAL SEARCH REPORT

Internati nal application N .
PCT/US94/03338

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of th relevant passages	Relevant to claim No.
X — Y Y Y Y Y	<p>US, A, 4,609,627 (GOLDSTEIN) 02 September 1986, see entire document.</p> <p>Biochemical and Biophysical research Communications, Volume 163, No. 3, issued 29 September 1989, S. Tsuji et al., "Molecular Cloning of a Full Length cDNA for Human α-N-Acetylgalactosaminidase (α-Galactosidase B)", pages 1498-1504, see entire document.</p> <p>Journal of Biological Chemistry, Volume 265, No. 35, issued 15 December 1990, A.M. Wang et al., "Human α-N-Acetylgalactosaminidase-Molecular Cloning, Nucleotide Sequence, and Expression of a Full-length cDNA", pages 21859-21866, see entire document.</p> <p>WO, A, 92/07936 (DESNICK ET AL) 14 May 1992, see entire document.</p> <p>Archives of Biochemistry and Biophysics, Volume 280, No. 1, issued July 1990, F. Yagi et al., "Glycosidases of Erlich Ascites Tumor Cells and Ascitic Fluid-Purification and Substrate Specificity of α-N-Acetylgalctosaminidase and α-Galactosidase: Comparison With CoffeeBean α-Galactosidase", pages 61-67, see particularly pages 61-63.</p>	<p>1-5, 13-17</p> <p>6-12</p> <p>6-12</p> <p>6-12</p> <p>6-12</p> <p>11-12</p>